Università degli Studi di Padova

Dipartimento di Elettronica e Informatica

Scuola di Dottorato in Ingegneria dell'Informazione

Indirizzo in Scienza e Tecnologie dell'Informazione

XXII Ciclo

Localization and network management in radio and underwater networks

Tesi di: *Francesco Zorzi*

Supervisore:
Prof. Andrea Zanella

Direttore della Scuola:
Chiar.^{mo} Prof. Matteo Bertocco

Anno Accademico 2009/2010

Università degli Studi di Padova

Dipartimento di Elettronica e Informatica

Scuola di Dottorato in Ingegneria dell'Informazione

Indirizzo in Scienza e Tecnologie dell'Informazione

XXII Ciclo

Localization and network management in radio and underwater networks

Tesi di: *Francesco Zorzi*

Supervisore:
Prof. Andrea Zanella

Direttore della Scuola:
Chiar.^{mo} Prof. Matteo Bertocco

Anno Accademico 2009/2010

Contents

Abstract

Wireless Sensor Networks (WSNs) have been widely studied in the last decade because of the large set of applications that can be potentially enabled by this novel network paradigm, as environmental monitoring, home automation, localization and tracking of mobile users, seamless and ubiquitous data exchange. However, the severe constraints in terms of energy, processing, memory and reliability of the low–cost sensor devices left a number of unresolved problems, open to research.

This thesis tackles two of the most interesting problems concerning WSNs, namely localization and traffic management.

Accurate sensor localization is crucial for a wide variety of WSN applications and protocols, including monitoring, routing, scheduling, data fusion and so on.

Typically, localization algorithms are based on an infrastructure of nodes which are aware of their geographical positions, called *beacons*. These nodes broadcast their coordinates in order to let the other nodes in the network, referred to as *stray* nodes, infer their own position by means of some estimation technique. The topic has been widely investigated, both in simulation and, more recently, in experimental testbeds. Nonetheless, the performance obtained by most of the proposed algorithms is still unsatisfactory, in particular in indoor environments. Therefore, further research is needed.

In this thesis, the problem has been tackled from different perspectives, in order to gain a deep and clear understanding of the several facets that characterize this interesting domain. As a consequence, we collected a rather wide set of results that apparently may not seem to be strongly correlated, though they all fit within the same research project.

As a first step, we compare well known localization algorithms in indoor environments for static networks, based on a Received Signal Strength Indicator (RSSI) ranging estimation technique. Results shed light on the actual potentialities and limits of common localization algorithms in indoor environments in a real testbed. In particular, we observed the strong impact of the shadowing random attenuation of the power of the received signal on the performance of the localization algorithms. Therefore, research has been addressed to the reduction of the shadowing term in the RSSI measures. To this end, we investigated the effect of the carrier frequency and of the antenna anisotropy on the randomness of the shadowing

component, again using real measurements collected in different testbeds.

Successively, our attention turned to the problem of beacon positioning that we observed may strongly affect the localization performance. We analyzed the optimal beacons placement both using an exhaustive but very complex approach and a heuristic scheme that provides close to optimal solutions while maintaining a linear complexity with the number of beacons.

Looking at the localization performance separating the different effects of localization algorithms, channel parameters and beacon positioning, has been important to understand the contribution of each one of these aspects to the localization error and how much is it possible to improve the localization accuracy by means of a single technique, that usually is chosen depending on the particular scenario and available resources.

Then nodes mobility has been included into the framework.

We first considered an Autonomous Mobile Robot (AMR) that can interact with sensors, but is also capable to self–localize using onboard odometry. Leveraging on the complementarities of WSN and AMR, we considered the Simultaneous Localization and Mapping (SLAM) problem that consists in creating the map of an area without any a priori knowledge of the environment, while localizing the nodes in the WSN by combining the information provided by the WSN and by the AMR odometry.

Second, we considered a completely mobile wireless networks in which heterogeneous nodes with different self–localization capabilities can interact one another on an opportunistic basis, exchanging localization information with other nodes that occasionally happen to be in geographical proximity. The performance of this scheme has been analyzed through a mathematical framework. We considered a Maximum Likelihood (ML) approach, a Linear Matrix Inequality (LMI) system and a simple heuristic strategy to define opportunistic localization algorithms. The validity of the mathematical model has been confirmed through experimental measures. We considered two main settings, one in which a node can rely on a single opportunistic interaction and one in which multiple contacts can be set while the node remains in the same position. We analyzed the different techniques, finding that in the first scenario, if the channel is not very noisy and the self–localization of the cooperative node quite accurate, the heuristic algorithm performs very well, in some case slightly better than ML, while the ML approach is very robust and improves localization accuracy even in difficult scenarios. If multiple interactions are available, then the heuristic technique is quite poor and it is better to rely on the LMI technique. Moreover LMI is quite flexible, because it does not require an initial guess of the position by the stray node and it can be used both with and without ranging information.

The static scenario analysis was very useful to focus on the mobile localization. The design of the proposed algorithms and scenarios, the simulation parameters and limits, are strictly related to the knowledge of the channel behavior and localization performance gained from the previous studies. The opportunistic idea raises from two main considerations: the

limited accuracy of the beacon based RSSI localization in a real scenario and the quite good ranging accuracy of RSSI when limiting the distance.

Alongside the main research line concerning localization in WSN, during the Ph.D. we also investigated other research topics, namely traffic management and underwater sensor network, which are not directly related to the previous ones but still of great scientific and educational interest. The most significant achievements obtained in these areas have also been collected in the thesis for two reasons. First, these topics still belong to the context of wireless sensor networks, sharing some basic characteristics such as the assumptions of simplicity and energy constraints. Second, the excursion on different but correlated fields may potentially open new perspectives to well known problems, thus contributing to the innovation and the progress of the research.

As mentioned, the second problem addressed in this thesis regards the traffic management in WSN. Usually all the nodes in a WSN send packets to a common node, called sink. This traffic pattern, under a certain load, can lead to congestion problems, causing packet losses, high delays and waste of energy. The proposed solutions in literature usually aim at detecting the occurence of congestion by involving in this task many nodes, sometimes the entire network. We propose a different protocol, called Efficient Packet Converge Casting (EPC^2), that mitigates the congestion at the sink involving only a fixed number of nodes, namely the sink's neighbors.

Another scenario we look into was that of underwater wireless sensor networks (UWSNs), which enable a number of applications as for radio WSNs, fostering interest in this research field.

Similarly to radio networks, the energy efficiency remains a main issue. Nodes are powered by battery and it is very important to extend the network lifetime as much as possible. The different characteristics of the environment in which nodes are deployed raise new research challenges that require novel protocol design.

We addressed the energy efficiency problem in UWSN with two different approaches.

First, we investigated the effect of duty–cycle and node density on the energy consumption of the network, assuming that nodes can use different power levels to transmit. Second, we proposed a channel management scheme to optimize the energy consumption, considering the strong relationship between distance, frequency and channel attenuation. Both solutions are very simple and suitable for the low–complexity underwater sensor devices and do not need any central unit to coordinate, but they work asyncronously and distributely.

Sommario

Negli ultimi anni, le reti wireless di sensori (WSN) sono state molto studiate a causa delle numerose applicazioni in cui possono essere usate, come il monitoraggio ambientale, la domotica, la localizzazione e il tracking di utenti mobili.Le forti limitazioni dei nodi sensori in termini di energia, processamento, memoria e affidabilità, lasciano ancora aperti molti problemi per la ricerca.

Questa tesi affronta due problemi molto importanti relativi alle reti wireless di sensori: la localizzazione e la gestione del traffico.

Un'accurata localizzazione dei sensori è importante per molte applicazioni per WSN, come monitoraggio, routing, scheduling, data fusion e molte altre.

Tipicamente, gli algoritmi di localizzazione si basano su una infrastruttra di nodi, detti *nodi áncora* che conoscono la loro posizione geografica. Questi nodi trasmettono in broadcast le loro coordinate agli altri nodi della rete, che da queste informazioni ricavano la loro posizione tramite tecniche di stima. L'argomento è stato largamente studiato, sia con simulazioni sia, più recentemente, con testbed sperimentali. Ciononostante, l'accuratezza ottenuta dalla maggior parte degli algoritmi proposti è ancora insufficiente, soprattutto in ambienti interni. È quindi necessario cercare nuove metodologie e nuovi approcci.

In questa tesi, il problema è stato affrontato da diversi punti di vista, in modo da capire in maniera più chiara e accurata i diversi aspetti che lo caratterizzano. Come conseguenza, abbiamo raccolto una vasta quantità di dati che potrebbero apparire come non molto legati uno all'altro, ma che in realtà rientrano tutti nello stesso progetto di ricerca.

Come primo passo, abbiamo confrontato algoritmi di localizzazione proposti in letteratura in uno scenario indoor e con nodi statici, stimando la distanza tra i nodi utilizzando la potenza del segnale ricevuto (RSSI). I risultati ci hanno permesso di capire le potenzialità e i limiti dei più diffusi algoritmi di localizzazione in ambiente indoor e in un testbed reale. In particolare, abbiamo osservato il grande impatto che ha sulle prestazioni di localizzazione l'aleatorietà, data dal termine di shadowing, della misura di potenza ricevuta. Abbiamo quindi cercato delle strategie per ridurre la varianza di questo termine aleatorio. A questo scopo, abbiamo studiato l'effetto della frequenza della portante, utilizzando una stima della potenza ricevuta multi–canale, e l'impatto dell'anisotropia dell'antenna sulle oscillazioni dei valori di potenza

ricevuta. Entrambi glil studi sono stati fatti con misure reali raccolte in diversi testbed.

Successivamente, abbiamo analizzato il problema del posizionamento dei nodi áncora, dopo aver osservato l'incidenza che questo ha sull'accuratezza della localizzazione. Abbiamo confrontato il posizionamento ottimo dei nodi áncora usando sia una tecnica esaustiva, ma computazionalmente molto complessa, sia uno schema euristico che raggiunge prestazioni molto vicine all'ottimo pur mantenendo una complessita' lineare con il numero di áncore.

Guardare alle prestazioni di localizzazione separando i diversi effetti degli algoritmi, dei parametri di canale e del posizionamento dei nodi áncora è stato importante per capire il contributo dei diversi aspetti all'interno dell'errore di localizzazione e quanto sia possibile migliorare la precisione della localizzazione ottimizzando uno di questi aspetti, che solitamente viene scelto in base allo scenario e alle risorse disponibili.

Quindi, abbiamo incluso nel nostro scenario anche nodi mobili.

All'inizio abbiamo considerato un robot mobile (AMR) che poteva interagire con i sensori, ma anche capace di localizzarsi grazie all'odometria. Facendo leva sulla complementarietà della rete di sensori e del robot mobile, abbiamo studiato e implementato un algoritmo di localizzazione e mappatura simultanea (SLAM), problema che consiste nel creare la mappa di un'area senza nessuna conoscenza a priori dell'ambiente e in contemporanea localizzare i nodi sensore confrontando le informazioni provenienti dai sensori e quelle ricavate dall'odometria del robot.

Poi abbiamo considerato uno scenario piu' generale composto da nodi mobili ed eterogenei, con diverse capacità di autolocalizzazione, che possono interagire uno con l'altro in modo opportunistico, scambiandosi informazioni di localizzazione con altri nodi che occasionalmente si trovano in prossimità. Le prestazioni di questo schema sono state analizzate in un modello matematico. Abbiamo studiato un approccio a Massima Verosimiglianza (ML), uno basato su Linear Matrix Inequalities (LMI) e una semplice strategia euristica per definire gli algoritmi di localizzazione opportunistica. La validità del modello matematico è stata confermata attraverso misure sperimentali. Abbiamo considerato due scenari principali, uno in cui un nodo può contare su una sola interazione opportunistica e uno dove possono essere fatti contatti multipli mentre il nodo resta nella stessa posizione. Abbiamo analizzato le diverse tecniche, trovando che nel primo caso, se le informazioni di autolocalizzazione del nodo cooperatore e di ranging sono buone, l'algoritmo euristico ha buone prestazioni, a volte addirittura meglio della Massima Verosimiglianza, che invece è estremamente robusto e riesce a migliorare la stima di localizzazione anche in scenari molto difficili. Se invece sono disponibili numerose interazioni, allora l'algoritmo euristico porta prestazioni scarse ed e' meglio utilizzare la tecnica LMI, specialmente utilizzando l'informazione di ranging. Inoltre l'LMI non richiede una conoscenza della posizione iniziale del nodo incognito.

Lo scenario statico è stato molto utile per studiare in maniera efficace la localizzazione mobile. La scelta degli algoritmi proposti e dello scenario, i parametri di simulazione e i

limiti, sono strettamente legati a quello che abbiamo studiato riguardo al canale wireless e alle prestazioni di localizzazione nei lavori precedenti. L'idea dello scenario opportunistico infatti e' venuta a partire da due considerazioni: la limitata precisione della localizzazione con áncore basata su RSSI in uno scenario reale e la buona precisione nella stima di distanza con RSSI quando la distanza è limitata.

Affianco al principale filone di ricerca riguardante la localizzazione nelle WSN, durante il dottorato di ricerca abbiamo approfondito anche altri argomenti, come la gestione del traffico e le reti di sensori sottomarine, che non sono direttamente collegate con il tema principale, ma sono comunque di grande interesse scientifico. I risultati più significativi ottenuti in questi temi sono stati inseriti all'interno della tesi per due motivi. Innanzitutto, questi argomenti appartengono al contesto delle reti di sensori wireless, condividendo alcune caratteristiche di base quali l'assunzione di semplicità e le limitazioni energetiche. Inoltre, il trattare campi diversi ma correlati, può aprire nuove prospettive a problemi noti, contribuendo così all'innovazione della ricerca.

Il secondo problema affrontato in questa tesi è stato la gestione del traffico in reti di sensori wireless. Spesso, i nodi di una rete di sensori mandano i pacchetti ad un nodo comune, chiamato sink. Questo modello di traffico, quando il carico cresce, può portare a problemi di congestione, causando perdita di pacchetti, ritardi e spreco di energia. Le soluzioni proposte in letteratura solitamente cercando di individuare l'inizio di una congestione, utilizzando in questo compito molti nodi, talvolta l'intera rete. Il protocollo proposto, chiamato Efficient Packet Converge Casting (EPC^2), mitiga la congestione al sink, ma coinvolgnedo solo un numero fissato di nodi, i vicini del sink.

Un altro scenario che abbiamo analizzato in questa tesi, sono state le reti sottomarine di sensori che, come nel caso delle reti radio, possono essere utilizzato per molteplici applicazioni e quindi hanno ricevuto molta attenzione dal mondo della ricerca.

Similmente alle reti radio, l'efficienza energetica è un problema molto sentito. I nodi sono alimentati a batteria ed è molto importante incrementare la vita della rete il più possibile. La profonda diversità dell'ambiente in cui i nodi sono disposti crea nuove sfide per la ricerca che richiedono la progettazione di nuovi protocolli.

Abbiamo affrontato il problema dell'efficienza energetica in reti sottomarine con due diversi approcci.

Abbiamo studiato l'effetto del duty–cycle and della densità dei nodi sul consumo energetico della rete, assumendo che i nodi potessero usare diversi livelli di potenza in trasmissione. Quindi abbiamo proposto uno schema di utilizzazione della banda disponibile per ottimizzare il consumo energetico, facendo leva sulla forte relazione tra distanza, frequenza e attenuazione del canale. Entrambe le soluzioni sono molto semplici e adatte ai dispositivi sottomarini che hanno forti limitazioni. Inoltre non richiedono una unità centrale per essere coordinate, ma operano in modo asincrono e distribuito.

Chapter 1

Introduction

In the last decade, a great interest has been raised by WSNs because of the large set of potential applications that can be enabled by this novel network paradigm.

The primary and most common application for WSN is environmental monitoring. The scenario consists in a large number of sensors distributed in a certain area that collect environmental information to be delivered and processed by a central unit, called *sink*. Other applications are home automation whose aim is to optimize the energy consumption in houses; localization and tracking of mobile users, as for example people, goods or vehicles in a battlefield or in a warehouse; seamless and ubiquitous data exchange, where the low–power and low–cost characteristics of WSN technologies enable the integration of the radio transceiver in a large number of portable devices. In the recent years, WSN applications have been proposed also in the underwater scenario, in particular for monitoring or environmental purposes.

Despite the effort spent to make WSN technology appetible for the commercial market, a number of still unresolved problems have slowed down the permeation of WSNs in the everyday life. In particular the severe constraints in terms of energy, processing, memory and reliability of the low–cost sensor devices pose challenges to researchers.

This thesis tackles two of the most interesting problems concerning WSN, namely localization and traffic management.

Accurate sensor localization is crucial for a wide variety of WSN applications. For example, in a monitoring system, it is important to locate the source of the information, in order to map the sensing measures in the area. Also protocol design can benefit from localization information, leading to more efficient solutions. Examples are the geographical routing algorithms [1, 2], which is proved to achieve very good performance by using nodes positioning information, or scheduling algorithms that take advantage of the location by optimizing the spatial reuse of transmission resources.

The topic has been widely investigated, both in simulation and, more recently, in experimental testbeds. Nonetheless, the performance obtained by most of the proposed algorithms

for WSN localization is still unsatisfactory, in particular in indoor environments. Therefore, further research is needed.

Typically, localization algorithms entail the presence of a limited number of nodes, called *beacons* (also known as *anchors* or *landmarks*), which are aware of their geographical positions. Beacons are required to periodically broadcast their coordinates in order to let the other nodes in the network, referred to as *stray* nodes, infer their own position by means of some estimation technique.

Localization algorithms can belong to two main classes: range–based, in which the stray node collects some distance estimates from physical measurements, e.g. received power, and, using some estimation technique, infers its position; and range–free, in which the stray node usually compares the measurements to estimate its position. A common physical measure is the Received Signal Strength Indication (RSSI), because every radio device natively has a circuit to trace this value and no additional hardware is required. Another ranging technique consists in the measure of the Time of Arrival of a signal, but it is not feasible in most off–the–shelf sensor devices, although it provides a good benchmark for comparison.

Usually, range–based algorithms are more accurate, but require the knowledge of channel parameters to estimate the distance from RSSI or ToA measurements, while range–free algorithms provide less accurate position estimate, but do not need to know the channel.

In order to gain a deep understanding of the localization problem, we followed a multistep approach. We firstly studied the localization problem in static networks to clearly comprehend the effect of very important network settings, like number of beacons, channel paramenters and beacons placement, on the localization accuracy. Successively, we considered the mobile scenario. This approach was very effective, because we have been able to focus on the problem exploiting the acquired knowledge. As a consequence, we present a rather wide set of results that apparently may not seem to be strongly correlated, though they all fit within the same research project.

As a first step, we conducted an accurate performance comparison among well known RSSI–based localization algorithms in indoor environments for static networks. The study was aimed at gaining better understanding of the actual potentialities and limits of common localization algorithms in indoor environments. Results have been obtained using measurements from real testbed. This study revealed the strong impact of the shadowing random attenuation on the performance of the localization algorithms as presented in [3].

Therefore, in order to ameliorate the localization performance, we studied some techniques to reduce the randomness of the received signal power. In particular we investigated the effect of channels hopping on the variance of the shadowing component and, consequently, on the accuracy of the localization estimation. Moreover we analyzed the effect of the anisotropy of the sensors antenna on the variability of the RSSI measurements.

We observed that the antenna anisotropy is responsible for a large fraction of the propa-

gation law variance and that channel hopping is also beneficial.

Successively, we addressed the problem of beacon positioning for effective nodes' localization. We observed that the location of beacons may strongly affect the localization performance and we analyzed the optimal setting both using an exhaustive approach whose complexity grows exponentially with the number of beacons and a heuristic scheme that provides a good beacon positioning while maintaining a linear complexity with the number of beacons.

In the following stage of our research work, we included nodes mobility in the framework.

We first introduced in the WSN an Autonomous Mobile Robot (AMR) that moves through the network and has its own localization information from onboard odometry, but can also communicate with the sensors.

Leveraging on the complementarities of WSN and AMR, we developed strategies to enhance the performance of both systems and enable novel functionalities. In particular we considered Simultaneous Localization and Mapping (SLAM) problem that consists in creating the map of an area without any *a priori* knowledge but only comparing the information provided by the WSN on the one hand and by the odometry on the other hand. Results have been collected in [4]

Finally we considered a completely mobile wireless networks in which heterogeneous nodes with different localization capabilities can interact on an opportunistic basis. Nodes are here supposed to be capable of self–localization, which can be performed by using some external infrastructure (e.g. beacon–based system, GPS) or some internal hardware (e.g. INS, MEMS, odometry). In addition, they can exchange localization information with other nodes that occasionally happen to be in proximity to improve their position estimate. A mathematical framework has been developed in [5] to study the performance of this scheme, using both a Maximum Likelihood approach and a very simple heuristic algorithm. Experimental data have been collected to confirm the validity of the mathematical model [6]. In this scenario we considered that a node can have only one contact with an opportunistic node. We also took into account a scenario in which a node can receive multiple responses to his opportunistic request, while staying in the same position. In this case, we proposed a localization algorithm based on the Linear Matrix Inequalities system to infer the position of a node completely unaware of his location [7, 8, 9, 10].

The second problem addressed in this thesis regards the traffic management in WSN. Usually all the nodes in a WSN send packets to the sink. Because all the traffic flows converge to a single node, in presence of a certain traffic load the network can suffer congestion problem losing reliability, due to packet losses and high delay, and wasting energy. Typical solutions introduce some end–to–end control to detect congestion. In certain solutions the entire network is involved in this task, even though the congestion is limited to a part of the network. Instead, in [11], we propose a protocol, called Efficient Packet Converge Casting

(EPC^2), that mitigates the congestion at the sink improving the management of the traffic near the sink. Furthermore EPC^2 involves only a fixed number of nodes, namely the sink's neighbors.

Another scenario we look into is the underwater wireless sensor networks (UWSNs). The activity on this new branch of WSN research has been growing very quickly in the last years, fostered by the wide range of potential applications it enables in monitoring, localization and tracking. Similarly to radio networks, the energy efficiency remains one of the main issue. Even in this scenario, nodes are powered by battery and it is very important to extend the network lifetime as much as possible. The different characteristics of the environment in which nodes are deployed raise new research challenges. First of all, nodes use acoustic communication, because the attenuation of electromagnetic waves in the water is too high. Acoustic waves, however, undergo very large propagation delays due to the speed of sound that is much slower than the light speed. Furthermore, the bandwidth is very limited, in the order of tens kHz. In addition, the signal strength attenuation depends more tightly on the distance than in radio networks. These features require new protocols for MAC and routing that take into account these new constraints while pursuing energy efficiency. We addressed the energy efficiency problem with two different approaches.

First, in [12], we investigated the effect of duty–cycle and node density on the energy consumption of the network, assuming that nodes can use different power levels to transmit. Second, in [13], we proposed a scheme of channel management to optimize the energy used in the transmissions, taking into account the strong relationship between distance, frequency and channel attenuation.

The thesis is organized as follows. In Chapter 2 we present an overview on the existing channel models. Chapter 3 describes the static localization problem, while in Chapter 4 opportunistic localization is presented. Chapter 5 deals with congestion at the sink and presents the EPC^2 algorithm. Energy efficiency in UWSNs is investigated in Chapter 6. Finally conclusions are drawn in Chapter 7.

Chapter 2

Experimental Characterization of indoor radio channel

An important problem when dealing with wireless networks is to model the radio channel response. In fact, the performance of radio communication and, more specifically, of the distance estimate methods based on the Received Signal Strength Indicator (RSSI), are strongly affected by the signal propagation conditions. Moreover, in order to simulate a WSN, an accurate statistical model of the channel is needed to provide realistic results.

In our work, in the specific, we need to estimate the distance between transmitter and receiver from some channel measurements, e.g. Time of Arrival (ToA), Angle of Arrival (AoA), Received Signal Strength Indicator. To this end, we need to derive a stochastic model that relates the measured value with the distance between transmitter and receiver.

In this chapter we fit our experimental results in a well known and rather general channel model and we investigate how frequency hopping and antenna anisotropy can affect the performance of ranging.

2.1 State of the art

The wireless channel has been widely studied and many models have been proposed to characterize the propagation of an electromagnetic signal.

Different models try to represent the channel with different degrees of accuracy or to describe a specific scenario. In fact, there is a tradeoff between the accuracy of a model and its complexity, in terms of computation, environmental information and hardware required for the analysis. Therefore approximations are often necessary.

The highest detail in signal propagation modeling is given by the Maxwell equations with boundary conditions, which take into account reflections, scattering and diffraction caused by every element in the environment. Obviously this method is not practicable because of

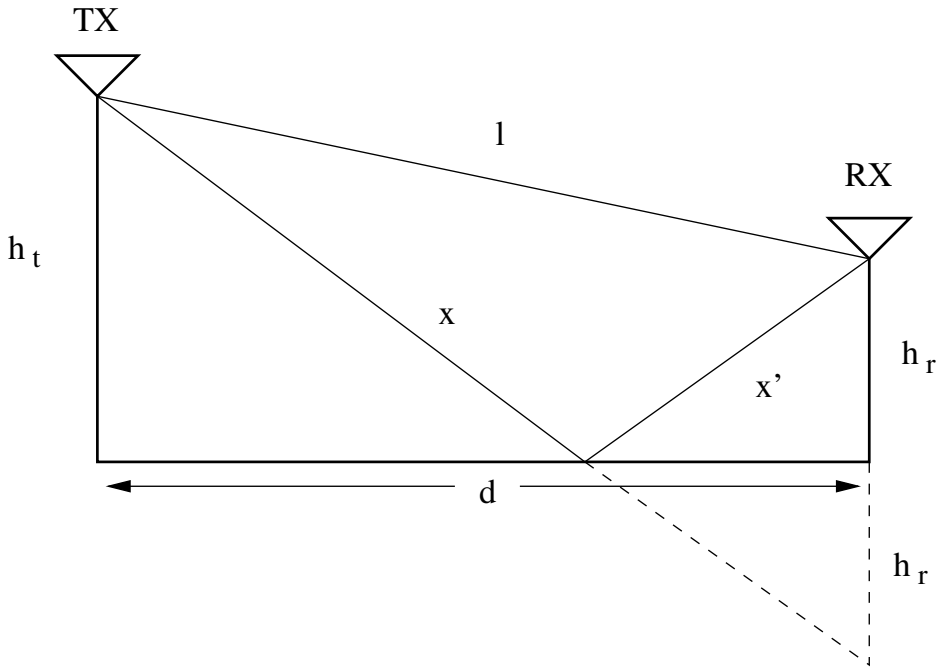


Figure 2.1: Example of two–ray propagation model

the very high complexity and the difficulty of collecting the required information.

A very common approximate model, called *ray–tracing*, considers the signal as a finite number of rays reflecting on the ground or on the objects. The simplest form is the two–ray model that is quite accurate if the direct path between transmitter and receiver is dominant (Line–of–sight, LOS).

Other approximate models are the empirical path loss, used to describe some particular wireless environments such as large urban macrocells, urban microcells and indoor scenarios.

At the base of all these models is the Friis formula for the free-space path loss ($P_L(d)$) that defines the ratio between the received and the transmitted power as

$$P_L(d) = \frac{P_r}{P_t} = \left[\frac{\sqrt{G_l}\lambda}{4\pi d}\right]^2 . \tag{2.1}$$

where P_r is the received power, P_t is the transmission power, G_l is the product of the transmit and receiver antenna gain in the LOS direction, λ is the wavelength and d the distance between transmitter and receiver.

Applying this formula to the two–ray scenario shown in Fig. 2.1, we can obtain the following relationship

$$P_r = P_t \left[\frac{\lambda}{4\pi}\right]^2 \left|\frac{\sqrt{G_l}}{l} + \frac{R\sqrt{G_r}e^{-j\Delta\phi}}{x + x'}\right|^2 \tag{2.2}$$

where R is the ground reflection coefficient, G_r is the antenna gain for the reflected path and

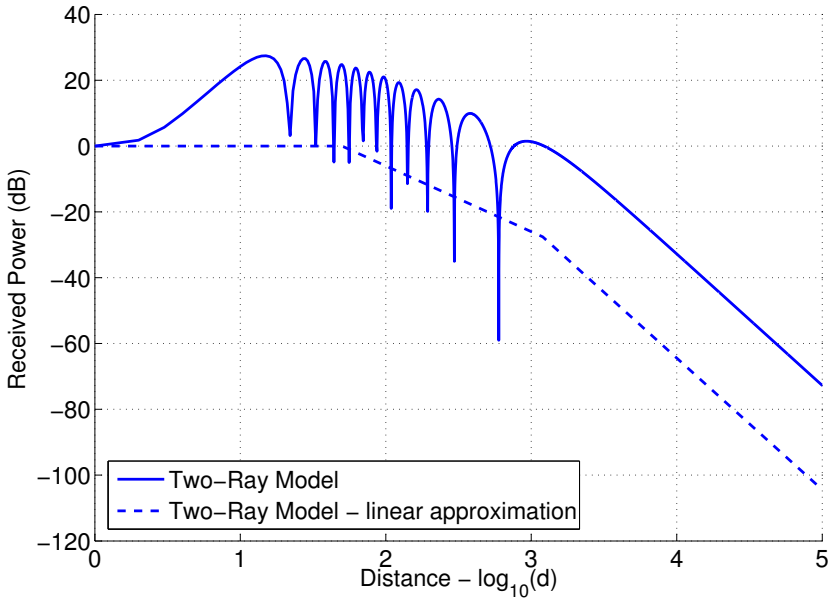


Figure 2.2: Received power varying the distance

$\Delta\phi$ is the phase difference between the LOS and the reflected paths.

When d increases, asymptotically we have $x + x' \approx l \approx d$, $G_l \approx G_r$ and $R \approx -1$ and the previous formula can be written as

$$P_r \text{ dBm } = P_t \text{ dBm } + 10\log_{10}(G_l) + 20\log_{10}(h_t h_r) - 40\log_{10}(d). \tag{2.3}$$

The received power (2.3) is shown in Fig. 2.2, where it is possible to recognize two thresholds at $d = h_t$ and $d = d_c = 4h_t h_r/\lambda$. For $d < h_t$ the power attenuation is constant, for $h_t < d < d_c$ it decreases at 20 dB/decade and for $d > d_c$ it decreases at 40 dB/decade.

This model is not very suitable for indoor scenario, where multipath interference is very strong, so that taking into account only one reflection is not so accurate.

The Okumura model tries to predict the signal propagation in large urban macrocells in the frequency range 150-1500 MHz. The formula is parameterized only in the distance and the central frequency (f_c), and the path loss attenuation is given by

$$P_L(d) \text{ dB } = L(f_c, d) + A_{mu}(f_c, d) - G(h_t) - G(h_r) - G_{AREA}, \tag{2.4}$$

where $L(f_c, d)$ is free space path loss, $A_{mu}(f_c, d)$ and G_{AREA} are additional attenuations due to the environment, $G(h_t)$ and $G(h_r)$ are the antenna height gain factor for the base station and the user respectively. Except for the path loss attenuation, the other prameters have been derived empirically from Okumura experiments [14].

Hata Model is a generalization of Okumura expressed in closed form. It refers to the same frequency range of Okumura model. The formula for urban scenario is given by

$$\begin{aligned} P_{L,urban}(d)\ \text{dB} &= 69.55 + 26.16\log_{10}(f_c) - 13.82\log_{10}(h_t) + \\ &\quad -a(h_r) + (44.9 - 6.55\log_{10}(h_t))\log_{10}(d), \end{aligned} \tag{2.5}$$

where $a(h_r)$ is a correction factor for the user antenna height based on the coverage area [14]. This model works quite accurately for distances larger than 1 Km, so that it is suitable for old cellular systems where cells were quite large, but not in the modern microcell systems that also use higher frequencies. Furthermore it models poorly indoor propagation.

Another approach is the Piecewise Linear Model or Multi–Slope, where the relation between attenuation and distance is described by a piecewise linear function, with N line segments on the distance axis. The breakpoints are chosen in many different ways, and best fitting slopes are estimated using a set of real measurements.

All these models are deterministic: once parameters are set, the attenuation value is fixed. This approach is not very suitable for indoor scenario, where moving objects and people create time–varying reflections, rifractions and scattering of the radio signal.

A very simple model for indoor environment consists of a deterministic path loss model to which some random effects are added, usually taking into account two different time-scale: one for fast but not very large variations, addressed as fast fading, and one for long term variation, addressed as slow fading or shadowing.

Most of the localization algorithms proposed in literature make use of the Received Signal Strength Indication (the RSSI) to get an estimate of the distance between transmitter and receiver (ranging). Unfortunately, the radio channel is highly unpredictable.

A possible way to dramatically improve the ranging accuracy consists in using other physical-layer measurements than RSSI, such as the time–of–flight of pressure waves, or ancillary radio hardware, such as multiple and/or directional antennas. However, these solutions are generally more energy demanding and/or require dedicated hardware, which results in more expensive devices. On the contrary, pure RSSI methods can be readily deployed in every wireless sensor network platform, since the RSSI circuitry is natively supported by most of the existing transceiver chipsets, with no extra hardware costs. Therefore, from the localization perspective, it is very important to have a good radio channel model to improve localization performance.

2.2 Radio channel modeling

We consider the rather general channel model proposed in [15]. According to this model, the received power $P_r(d,t)$ from a node in position $\mathbf{s} = [x, y]$ at a distance d from the transmitter

at time t can be expressed by the following equation [in dB]:

$$P_r(d,t) = P_t + K + 10\eta \log_{10}\left(\frac{d}{d_0}\right) + \alpha(t) + \psi(\mathbf{s},t) \tag{2.6}$$

where P_t is the transmission power [in dB], K is a unitless constant that depends on the environment, d_0 is the reference distance for the antenna far field, and η is the so–called path loss coefficient that depends on the environment. The term $\alpha(t)$ is a Rayleigh–distributed random variable that takes into account the fast fading effect, while $\psi(\mathbf{s},t)$ is a Normal random variable, with zero mean and variance σ^2_ψ, that describes the shadowing effect. The $\mathbf{s}$ dependancy means that the value of ψ is not related to the distance, but more specifically on the relative positions of transmitter and receiver. It is worth to notice that some of these parameters depend on both time and space, while others depend only on space and some of them are independent of both. For example, the transmission power is not related to the particular scenario in which nodes are involved, whereas the fast fading component changes very quickly in time, almost independently on the location. Moreover, it is important to underline that Rayleigh random variable has zero mean. Assuming that different realizations of $\alpha(t)$ are independent and identical distributed, it is possible to average out the fast fading term in (2.6) by taking multiple measures in different time instants. Unfortunately $\psi(\mathbf{s},t)$ varies both in space and time, even though usually the value changes more rapidly in space than in time.

Therefore, averaging over multiple measurements between the same couple of nodes, it is possible to approximate (2.6) by neglecting the fast fading term (2.7) and assuming the shadowing depends only in space,

$$P_r(d) = P_L(d) + \psi(x) \tag{2.7}$$

where

$$P_L(d) = P_t + K + 10\eta \log_{10}\left(\frac{d}{d_0}\right) \tag{2.8}$$

In our study, we give particular relevance to the relation between received power and distance. We can disclose right now that ranging estimate is more accurate when σ_ψ, i.e. when the variability of the received power is low, and when nodes are close, because the estimation error is proportional to the distance between transmitter and receiver.

Using (2.8), it is possible to infer an estimate of the distance given the received power as

$$\hat{d}^{rssi} = d_0 10^{\left(\frac{P_t+K+\psi(x)-P_r(d)}{10\eta}\right)} = d \cdot 10^{\frac{\psi(\mathbf{s})}{10\eta}} \tag{2.9}$$

It is possible to notice that this distance estimator is biased, i.e., the expectation of $\hat{d}$ is not the true value of d:

$$E[\hat{d}^{rssi}] = E[d \cdot 10^{\frac{\psi}{10\eta}}] = d \cdot E[e^{\frac{\psi \log 10}{10\eta}}]. \tag{2.10}$$

We know that ψ is a Gaussian random variable $N(0, \sigma_\psi)$, we can define $\psi' = \psi \log 10/(10\eta) \sim N(0, \sigma_\psi \log 10/(10\eta))$. Therefore $e^{\psi'}$ is log-normal and its expectation is given by

$$E[e^{\psi'}] = \exp\left(\mu_{\psi'} + \frac{\sigma^2_{\psi'}}{2}\right) = \exp\left(\left(\frac{\sigma_\psi \log 10}{10\eta}\right)^2 \frac{1}{2}\right) \tag{2.11}$$

Therefore, for a given distance d, the mean m_{rssi} and variance σ^2_{rssi} of the ranging error $e_{rssi} = \hat{d}^{rssi} - d = d(1 - e^{\psi'})$ are given by

$$m_{rssi} = d\left(1 - e^{\sigma^2_{\psi'}/2}\right) \tag{2.12}$$

$$\sigma^2_{rssi} = d^2\left(e^{\sigma^2_{\psi'}} - 1\right) e^{\sigma^2_{\psi'}} \tag{2.13}$$

From (2.13) we can note that the ranging error is multiplicative with the distance

We now describe the Time of Arrival (ToA) ranging model. This scheme usually achieves better performance than RSSI but it requires a more expensive hardware.

In the ToA method, the distance estimate is performed by measuring the time of flight of radio signals exchanged by node and beacon. According to [16], the relation between time of flight and distance is given by

$$T(d) = \frac{d}{v_p} + \mu_T + \psi_T(\mathbf{s}) \tag{2.14}$$

where $T(d)$ is the time that a signal takes to propagate from the transmitter to the stray node in position $\mathbf{s}$ at a distance d, v_p is the propagation speed of the signal, and $\mu_T \geq 0$ is a range bias induced by processing time and non line-of-sight propagation. Finally, $\psi_T(\mathbf{s})$ accounts for the stochastic measurement errors and it is generally modeled as a Gaussian random variable with zero mean and variance σ_T. The estimated distance $\hat{d}^{toa}$ is, hence, given by

$$\hat{d}^{toa} = (T(d) - \mu_T)v_p = d + \psi'_T(\mathbf{s}) \tag{2.15}$$

where $\psi'_T(\mathbf{s}) = \psi_T(\mathbf{s})v_p$. The mean m_{toa} and variance σ^2_{toa} of the ranging error at distance $e_{toa} = \hat{d}^{toa} - d = \psi'_T$ turn out to be

$$m_{toa} = 0 \tag{2.16}$$

$$\sigma^2_{toa} = v_p^2 \sigma_T^2 \tag{2.17}$$

Note that, in this case, the error is independent on the distance, so that the ranging error is identically distributed irrespective of the actual distance between node and beacon.

2.2.1 Channel parameters estimation

In order to use the model (2.7), it is necessary to estimate the channel parameters η and K. The estimation of the statistical distribution of $\psi(x)$ gives some information on the accuracy of each power measurements compared to the estimated model.

To estimate the parameters η and K, we used the Least Square approach. From a set of measurement pairs (p_i, d_i), with $i = 1, 2, \cdots, N$, the deterministic part of the propagation model can be written as

$$\mathbf{P} = \mathbf{X} \begin{bmatrix} K \\ \eta \end{bmatrix} \tag{2.18}$$

where

$$\mathbf{P} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_N \end{bmatrix} \tag{2.19}$$

and

$$\mathbf{X} = \begin{bmatrix} 1 & -10 \log \frac{d_1}{d_0} \\ 1 & -10 \log \frac{d_2}{d_0} \\ \vdots & \\ 1 & -10 \log \frac{d_N}{d_0} \end{bmatrix} \tag{2.20}$$

Using the LS method, we can find the estimated values of the two unknown parameters K, η as

$$\begin{bmatrix} \hat{K} \\ \hat{\eta} \end{bmatrix} = \left(\mathbf{X}'\mathbf{X}\right)^{-1} \mathbf{X}'\mathbf{P}. \tag{2.21}$$

After that, it is possible to estimate the values of the realizations of the shadowing random variable as

$$\hat{\mathbf{\Psi}} = \mathbf{P} - \mathbf{X} \begin{bmatrix} \hat{K} \\ \hat{\eta} \end{bmatrix} \tag{2.22}$$

Then, the mean and the variance of ψ can be estimated as

$$\hat{m}_\psi = \frac{1}{N} \sum_{k=1}^{N} \hat{\psi}_i \tag{2.23}$$

$$\hat{\sigma}^2_\psi = \frac{1}{N-1} \sum_{k=1}^{N} (\hat{\psi}_i - \hat{m}_\psi)^2 \tag{2.24}$$

2.3 Experimental Testbed

In our study, we performed experiments in different testbeds and with different devices, in order to achieve a better understanding of the channel behaviour.

We used IFX-Eyes sensor nodes [17], which have been extensively used by our research group to test protocols and algorithms for classical Wireless Sensor Networks (WSN) [18, 19]. The nodes can be programmed and powered via USB, thus permitting easy interconnection

with other digital devices. Each board is equipped with a radio interface that provides 19.2 kbps transmission rate by using FSK modulation in the 868.3 MHz band. The platform is also equipped with light and temperature sensors. Furthermore, a Received Signal Strength circuit returns a two-byte integer proportional to the received signal strength.

TmoteSky sensor nodes [20] are equipped with an IEEE 802.15.4 compliant radio transceiver (PHY and MAC layers) operating in the 2.4 GHz ISM band with DSSS modulation and transmission rate of 250 kbps. The transceiver can scan the spectrum over 16 channels, spaced by 5 MHz, from 2.405 GHz to 2.480 GHz. The transceiver also provides the values of Received Signal Strength Indicator and the Link Quality Indicator (LQI).

2.3.1 SIGNET testbed

The first testbed consists of a room measuring approximately 10×10 meters in which we deployed 48 nodes, both EyesIFX and TmoteSKY, so that the distances among nodes vary from 1.1 to 12.5 meters. The nodes are placed on a grid and suspended at approximately 75 centimeters from the ceiling. The testbed layout is shown in Fig. 2.3. Note that the presence of moving people causes the radio channel to be time varying and highly affected by multipath interference.

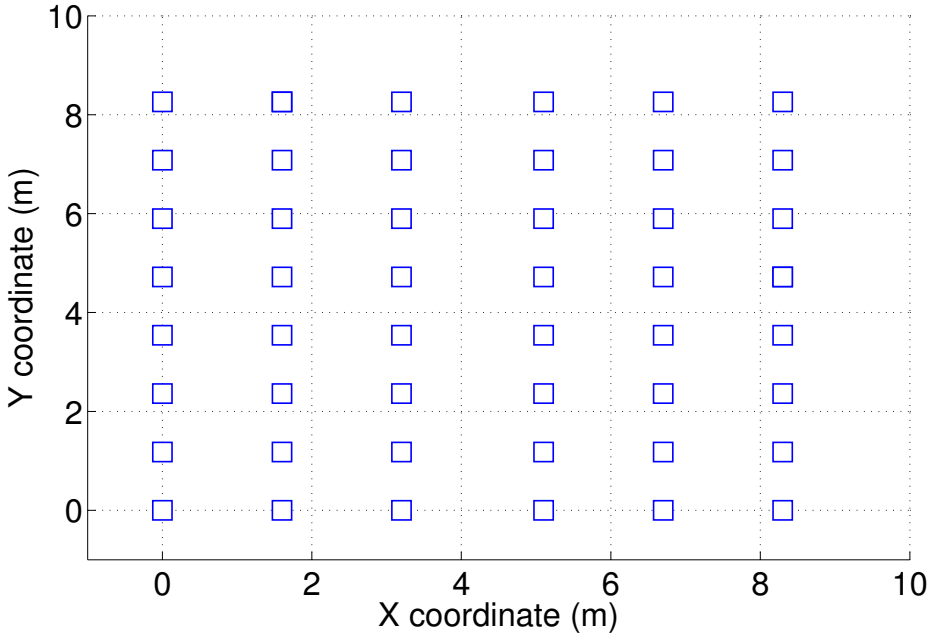


Figure 2.3: SIGNET testbed

2.3.2 Mobile testbed

The second testbed consists of a corridor where 10 EyesIFX sensor nodes and a mobile node are deployed, as shown in Fig. 2.4. The mobile node moves along a preplanned path drawn in solid line in the figure and ha the capability to know its own position. In this case, transmissions occur only among the mobile node and the static nodes. The distance between transmitter and receiver varies from few centimeters to almost eight meters. In this scenario transmissions have been performed in a single RF channel.

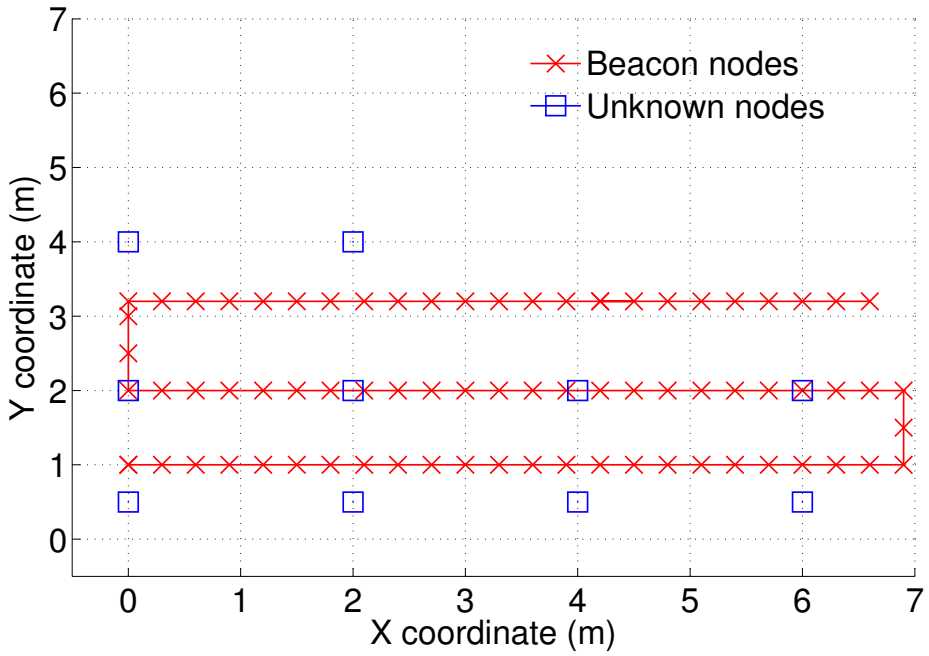


Figure 2.4: Mobile testbed

The robot

The robots is a wheeled custom designed robot based on the Pioneer 2 by MobileRobots Inc. The robot is equipped with a standard ATX motherboard with a 1,6 GHz Intel Pentium 4 processor, a 256 MB RAM memory and a 160 GB hard disk storage, running Linux OS. The only on–board sensors are an omnidirectional camera (composed of a standard CCD camera and a convex omnidirectional mirror) and the odometers connected to the two driven wheels. Communication is provided by a PCMCIA wireless ethernet card and by an EyesIFX node connected to one of the robot's USB ports. The EyesIFX sensor is mounted on top of the omnidirectional camera.

Mounting an EyesIFX node on the robot enables to integrate the robot in WSN. Thus the robot itself can act as a gateway for the WSN services. To seamless use the special command

set of the EyesIFX mounted on the robot from the robot control software, we exploited the robotics architecture MIRO [21, 22].

Miro is a distributed object oriented framework for mobile robot control, based on CORBA[1] technology. The Miro core components have been developed in C++ for Linux, but due to the programming language independency of CORBA further components can be written in any language and on any platform that provides CORBA implementations. The Miro core components have been developed under the umbrella of ACE[2], an object oriented multi-platform framework for OS-independent interprocess, network and real time communication. They use TAO[3] as their ORB[4]. TAO is a CORBA implementation designed for high performance and real time applications.

Miro EyesIFX service

Miro supports several robot sensors such as: lasers, kickers, sonars and several camera types. Miro also supports the most used wheeled robot platforms like: Pioneer bases, Sparrow bases and B21 bases. Unfortunately, there are no MIRO modules to support communication with a WSN. Thus, we wrote a new Miro service to allow bidirectional communication between the robot (or potentially, the robots) and the WSN, by extending the `Server` Miro class. The new Miro service is called `Eyes Service`. We created an Eyes Message object as an unique structure to send or receive data between the WSN and the robot. The message is an ordered buffer filled with the command to be send (or received), the message options and the data.

Conversely, the communication process is managed by two different classes: the `EyesConnection class` and the `EyesEvent class`. The first class takes care of the communication from robot to sensor and therefore to WSN. In this class we initialize the sensor port, reset the sensor, verify if it works and implement the send primitive. The second class takes care of the communication *from sensor to robot*. This class is an object wrapper for the event handler registered with the ACE ReactorTask. From ACE event handler, this class inherits the function that handles file descriptor event, that we overwrite to read from sensor serial port, create the Eyes Message and pass it to decoder task.

To use the services of this structure, we exploit TAO services, in particular the *Name Service*. When the *Name Server* is running, services can be shared easily adding them to a server or they can be requested to the Name Server. We use this service with the multicast support which hides the Name Server address. TAO implements several ways to share data, we decided to use the producer/consumer paradigm. TAO translate this in a supplier/consumer structure which needs a Notify Channel to share data.

[1] Common Object Request Broker Architecture
[2] Adaptive Communications Environment
[3] The ACE ORB
[4] Object Request Broker

2.3.3 Lecture Hall testbed

The third testbed consists of 25 TMoteSKY nodes deployed sparsely in the Lecture Hall of the Department of Information Engineering. Nodes transmit in all the RF channels provided by the TMoteSky devices. The nodes are positioned upon chairs, with distance between 1 and 14 m. Fig. 2.5 shows the disposition of the nodes in the testbed.

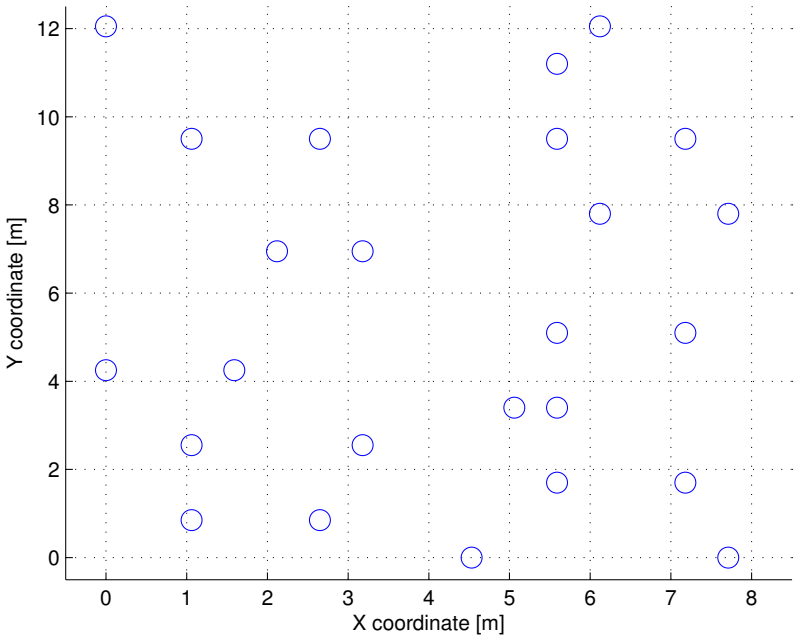


Figure 2.5: Lecture Hall testbed

2.3.4 Opportunistic testbed

The opportunistic testbed consists in a room measuring approximately 9 m $\times$ 11 m with 2 m high ceiling in the lowest area. The room is empty, except for the central support beam. TMoteSky Sensor nodes are deployed to form a grid, covering a rectangular area of size 7.2 m $\times$ 9 m, as shown by the blue circles in Fig. 2.6. Nodes are placed on the top of plastic cones 30 cm high from the floor. In addition to the 24 static nodes placed at the grid intersections, the testbed also includes 4 beacon nodes placed at the corner of the room, in the positions marked by red circles in Fig. 2.6. Beacons have perfect knowledge of their own position and can be used by the other nodes for (non-oppotunistic) self-localization, e.g., using the well-known Min-Max algorithm [23]. Finally, a so-called *opportunistic node* has been placed in 120 different positions along the path marked by the black line in Fig. 2.6.

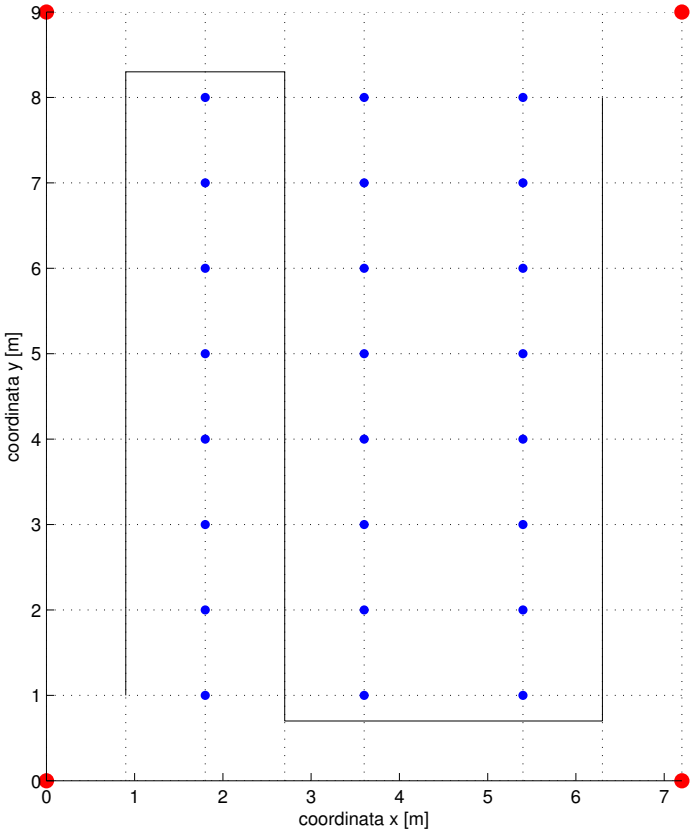


Figure 2.6: Testbed deployment topology

2.4 Experimental results

Most of the localization techniques in WSNs rely on RSSI measurement to infer the distance estimation. Therefore, it is important to have a realistic channel model and to develop some schemes to reduce the randomness of the channel measurements in order to have more reliable distance estimations. To this end, we collected a large number of measurements set from different testbeds and different sensor nodes, using both TMoteSky and EyesIFX nodes. When we used TMoteSky nodes, we exploited their capability of using 16 different carrier frequencies, spaced by 5 MHz. The objective of the study was to understand if channel parameters and distance estimation are affected by using a set of the available carrier frequencies. Moreover we studied the impact of the antenna propagation effect on the long term fading.

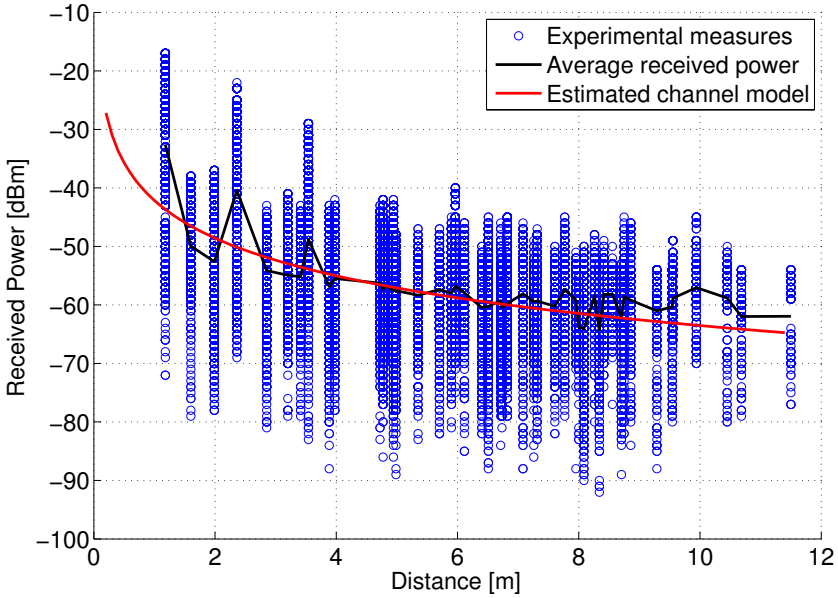


Figure 2.7: Experimental measurements - SIGNET testbed

	$\hat{K}$ (dB)	$\hat{\eta}$	$\hat{\sigma}_\psi$ (dB)
SIGNET testbed	23.18	2.25	7.34
Lecture Hall testbed	33.17	1.77	5.87
Mobile testbed	44.12	1.51	6.34
Opportunistic testbed	-36	1.74	6

Table 2.1: Channel estimation values

2.4.1 Single channel analysis

Fig. 2.7 shows the received power as a function of the distance between trasmitter and receiver. The circles represent the actual power measurements for that distance. As expected, the measured values have some variability due to fast fading. The average of these measures is represented by the black piecewise line in the graph. The smooth line follows the theoretical behavior given by (2.7) without considering the random term ψ and with parameters estimated as (2.21).

Then, we analyzed the statistical distribution of the shadowing, given by (2.22). In Fig. 2.8, we show the qq–plot of the shadowing, that is a comparison with a gaussian random variable with zero mean and variance $\hat{\sigma}_\psi$ as obtained from (2.24). The graph confirms that the hypothesis of the gaussian distribution for ψ is reasonable.

We repeated the analysis for the other scenarios described above. Tab. 2.1 summarizes the estimated values of the channel model parameters in the two cases.

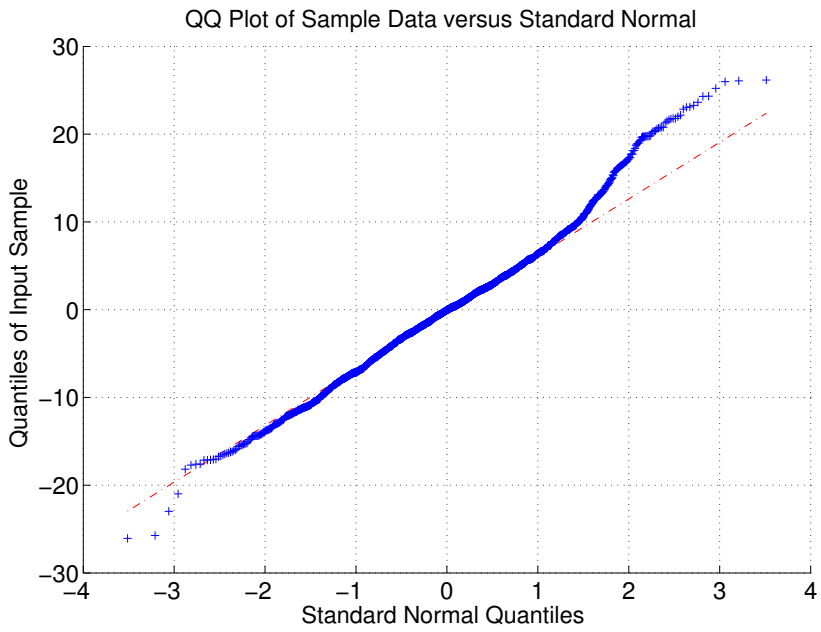


Figure 2.8: qq–plot of shadow fading - SIGNET testbed

It is interesting to see that $\hat{\eta}$ can have a value smaller than 2, which is the value in case of free space propagation. This is likely due to the fact that, in indoor scenarios, reflections may add constructivily to the line-of-sight propagation ray. In the Mobile testbed the value of $\hat{\eta}$ is even smaller because the place where measures have been collected is a corridor, which provides a sort of wave guide effect. In the SIGNET testbed the fast fading is higher due to the presence of moving people and objects. The statistical power of shadowing instead does not change very much among the different scenarios.

Looking at these results from a localization point of view, we realize that the distance estimation between transmitter and receiver devices through RSSI is very challenging. In fact, according to (2.10), the accuracy of the ranging estimate depends on the ratio between σ_ψ and η that it is quite high. The value of η depends on the environment and can not be varied. Conversely the variance of the shadowing may be reduced by taking advantage of the frequency hopping capability of the sensors as described in the next section.

2.4.2 Multiple channels analysis

To evaluate the impact of RF hopping on the channel parameters estiamte, we first compare the estimates obtained by considering each RF channel separately. Tab. 2.2 shows the results for the SIGNET testbed. It is worth to notice that the parameters obtained for different RF channels are indeed very similar, with $\hat{\eta}$ ranging between 2.15 and 2.37, while $\hat{\sigma}_\psi$ between 7.21 and 7.64 dB. The value of $\hat{K}$ shows slightly larger variations, ranging from 19.97 and 27.20

f (MHz)	2410	2420	2440	2450	2460	2470	2480	**ALL**
$\hat{\eta}$	2.177	2.148	2.306	2.299	2.238	2.327	2.371	**2.252**
$\hat{\sigma}^2_\psi$ (dB)	51.983	56.553	53.747	52.709	53.386	58.359	52.425	**35.354**

Table 2.2: Channel parameters estimated for different frequencies

Channel used (MHz)	$\hat{\sigma}^2_\psi$ (dB)
[2460 2465 2470]	48.35
[2440 2445 2450]	45.71
[2020 2050 2080]	37.57
[2040 2050 2060 2070]	40.44
[2040 2045 2050 2055 2060 2065 2070]	40.49
[2010 2020 2040 2050 2060 2070 2080]	35.37

Table 2.3: Variance of the estimation error using different subset of channels

dB, but it does not affect much the distance estimation because it is an additive contribution. Therefore, using a single RF channel, there is no reason to prefer one RF channel to another, because the signal behavior is very similar.

As a second step, we estimated the channel model parameters by considering simultaneously measures collected on different RF channels. In this case the power value at a given distance was obtained as the average of all measures collected for that distance over time and for a subset of C RF channels, i.e.

$$p_i = \frac{1}{NC} \sum_{n=1}^{N} \sum_{c=1}^{C} p_i^{n,c} \tag{2.25}$$

where $p_i^{n,c}$ denotes the n-th received power sample taken at RF channel c by pairs at distance d_i. Using the entire dataset, we obtained $\hat{\eta} = 2.25$, $\hat{K} = 23.18$, which are very similar to the values obtained for single channels. Conversely $\hat{\sigma}_\psi$ decreases to 5.93 dB, which means 20% smaller than the single channel scenario.

This improvement in terms of reduction of the variance of the power shadowing is rather important because it translates in better localization accuracy.

In addition, we investigated the effect of the number of RF channels considered for the parameters estimation on the variance of ψ. In fact, it is clear that if two channels are strongly correlated, the information gained using both channels is negligible compared to the single channel case.

We compared the results obtained considering several subsets of channels, that differ in the number C of channels considered and in the frequency gap between channels. Results shown in Tab. 2.3 confirm that neighbor channels are rather correlated. Moreover a significant reduction of σ_ψ can be obtained using a suitable selection of channels.

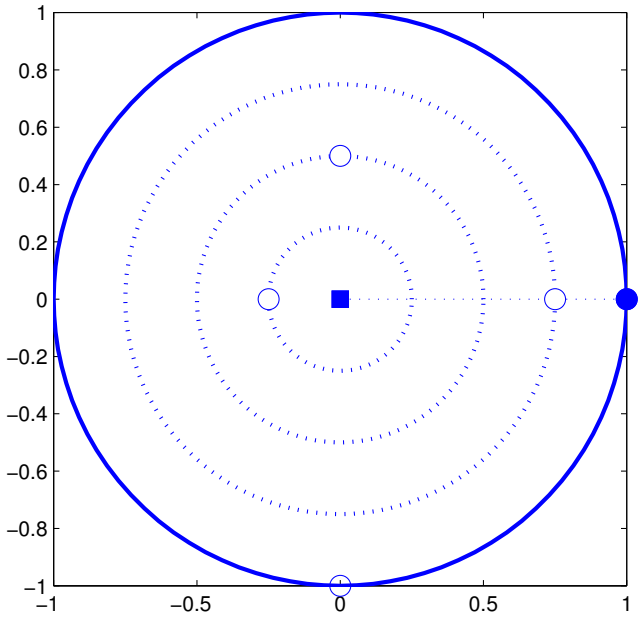


Figure 2.9: Experimental setup for the analysis of the antenna anisotropy effect

For instance, considering four channels separated by 10 MHz, the reduction of σ_ψ^2 is not as good as using three channels spaced by 30 MHz apart. Moreover adding three additional channels not sufficiently spaced from the others, we do not observe any enhancement. On the contrary, if we take seven channels spread on the entire bandwidth, then the reduction of σ_ψ^2 is exactly the same as using all the channels.

Therefore, good performance in terms of distance estimate is expected by using a multi-channel approach with few sufficiently spaced RF channels.

2.4.3 Antenna propagation effect

To better understand the nature of the random errors that affect the path loss model, we considered the effect of the antenna propagation diagram on the received power.

To this end, we set up an experiment as in Fig. 2.9. We used a fixed node (the square at the center of the graph) and another ndoe (the circle) that could be positioned at several distances between 20 cm and 2.4 m from the first one, with steps of 10 cm. The node was moved along a circular trajectory with rotation angle of 10 degrees. We collected several measurements for each position. The experiment was repeated with two modalities.

In the first one, the central node was fixed, so that the relative angle between the antennas of the two nodes changed everytime the rotating node moved circularly. In the second one, the central node was rotated in an integral manner with the external node. Therefore the two

nodes were always facing each other in the same way. Results in terms of RSSI vs distance are reported in Fig. 2.10 and Fig. 2.11, for the first and second experiment respectively. We clearly see how the anisotropic antenna radiaton pattern affects the received power at any distance.

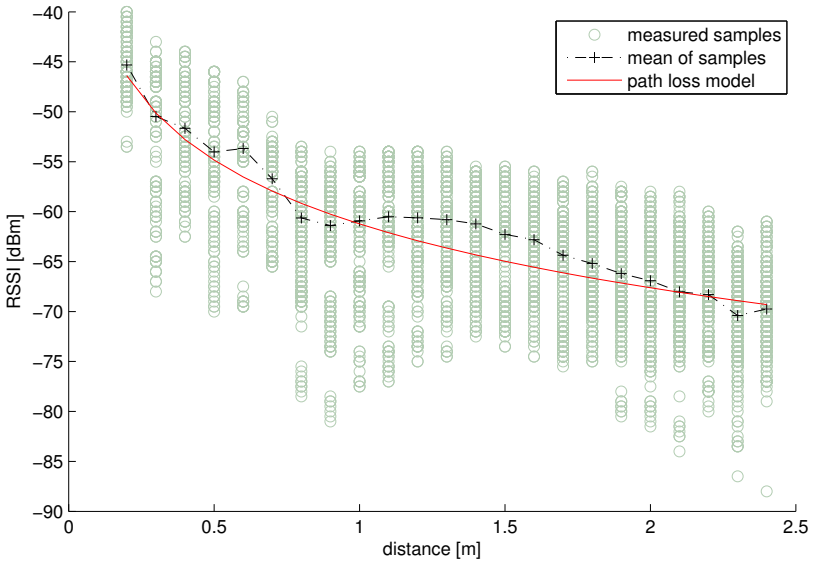


Figure 2.10: Received power vs distance with variable antennas orientation

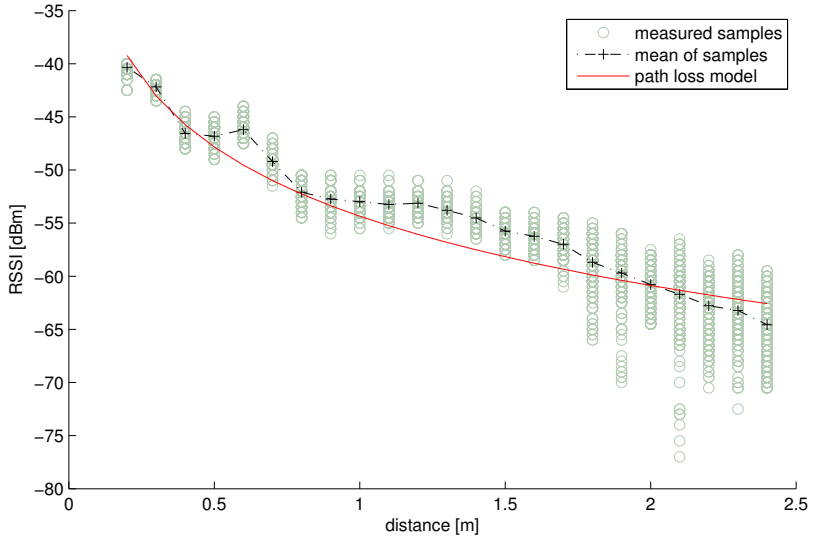


Figure 2.11: Received power vs distance with fixed antennas orientation

Tab. 2.4 summarizes the channel parameters estimated in the two experiments, given

$P_t = -1$ dBm and $d_0 = 10$ cm. We observe that the path loss parameters do not depend on the anisotropy of antenna. This result is coherent with the one in the previous section, where the path loss exponent did not change using a more complete information set obtained from the measures collected in different RF channels. The value of K changes due to the hardware offset introduced by the different gains of the antennas of the two nodes.

The most impressive difference is the reduction of the $\hat{\sigma}_\psi$ when the angle between the nodes is fixed, with respect to the case when the orientation of the two antennas is random. Looking at the graph in Fig. 2.11, we note that small distances are not affected by channel shadowing. In this situation, the reflections and refractions of objects are not so strong compared to the line–of–sight path, while increasing the distance, the behaviour is more unpredictable. The behaviour in Fig. 2.10 shows a much higher variance on the received power compared to the previous case, leading to a very inaccurate estimation of distance using the received power. This confirms that the antenna radiation pattern is a very important issue to be considered when using RSSI measurements for ranging purposes.

K (dB)	η	σ_ψ (dB)
-39	2.12	5.13
-31.7	2.16	2.2

Table 2.4: Channel Parameters' estimate with variable antennas orientation (first row) and fixed antennas orientation (second row).

Chapter 3

Localization in Wireless Sensor Networks

Most of the localization algorithms proposed in literature make use of the Received Signal Strength Indication (the RSSI) to get an estimate of the distance between transmitter and receiver (ranging).

The study of the channel in Chapter 2 is fundamental to approach the localization problem. Analyzing the collected channel measurements, we know that the radio channel is highly unpredictable and because localization algorithms are based on these measures, also localization performance suffers, in terms of accuracy, this variability. Moreover, using real data, we obtained realistic channel parameters to use when we simulated the localization algorithms. Therefore, these channel parameters were consistent with the real data, leading to realistic results.

In this chapter we investigate different aspects of the localization problem, taking into account the WSNs hardware limitations.

As a first step, we set up two different WSN testbeds in order to collect real measurements in an indoor environment using a static network. We used these data to compare experimental performance of well-known RSSI–based localization algorithms using beacon nodes. The aim of this work was to achieve a better understanding of the actual potentialities and limits of these localization algorithms.

Successively, we turned our attention to the beacons, investigating the effect of beacon positioning on nodes' localization performance in a static network, and considering both RSSI and Time of Arrival (ToA) ranging techniques.

Finally we introduced an autonomous mobile robot (AMR) into the static WSN scenario. The AMR is capable of performing good self–localization and tracking based on onboard odometry. Moreover we equipped the AMR with a sensor node, in order to enable communication with the WSN. We hence studied a Simultaneous Localization And Mapping (SLAM)

algorithm.

3.1 State of the art

As previously mentioned, the literature on this topic is huge and an accurate survey is clearly out of the scope of this section. Rather, we focus on RSSI based schemes only, dwelling upon the algorithms that have been used in our study. A more comprehensive overview of the topic can be found, for instance, in [24, 25, 16].

Range–based algorithms make use of the RSSI to estimate the distance between nodes. Then, different techniques, such as lateration [26], triangulation [23] or statistical inference [27], are used to estimate the position of stray nodes with respect to the beacons. Unfortunately, RSSI–based ranging is severely affected by errors due to the unpredictable radio propagation behavior, especially in indoor environments. A poor ranging usually determines very loose position estimates and, hence, unsatisfactory localization performance. This limit has been remarked in [28, 29], where the authors state that the performance cannot improve significantly even using complex algorithms, being intrinsically limited by the ranging errors. However, such papers do not investigate how the ranging errors are affected by the number of anchor nodes and the accuracy of the radio propagation model. The ranging errors due to noisy RSSI values, however, can be mitigated through a series of refinement phases, as in Savarese [30] and Savvides [23].

In RSSI–based range–free algorithms, localization is still performed by exploiting the RSSI values that, however, are not used to estimate the distance between stray and beacon nodes. Good results have been obtained by RSSI–mapping techniques, such as RADAR [31], which require to preliminary perform an accurate measurement campaign aimed at constructing a map of the radio signal strength received in the area of interest. Comparing the RSSI values received from the different beacons with the pre-built RSSI map, a node can estimate its own position in the area. Another family of range–free algorithms makes use of the RSSI samples to establish order relationships between nodes. A great advantage of this approach is its independence of the underlying channel. On the other hand, it makes it difficult to reach a low localization error even in the presence of a good radio channel.

In our experimental activity, we consider four algorithms, namely *Min–Max* [32, 33], *Multilateration* [32, 33] and *Maximum Likelihood* [27, 34] for the range–based category, and *ROCRSSI* [35] as a representative for the range–free algorithms.

In [36] the problem of beacon positioning is addressed. The authors proposed a technique to refine the beacon placement based on the collected measurements. In our work, we want to find a more general scheme, that can be valid for every scenario.

Localization problem becomes more complex when the position of the landmarks is also unknown, in which case we better speak of SLAM (Simultaneous Localization and Mapping).

Kurth, et al [37] have developed a SLAM system based on the time of flight of only radio signals. The achieved localization error is of the order of one meter.

3.2 Experimental localization results in indoor environment

In this section we address the localization problem by presenting an accurate performance comparison among some well known RSSI–based localization algorithms in indoor environments. The aim of the study is to achieve a better understanding of the actual potentialities and limits of common localization algorithms in indoor environments and to shed light on some still open issues, such as how many beacon nodes are needed to reduce the localization error below a certain limit, and whether complex estimation algorithms, such as Maximum Likelihood (ML), outperform simple ones, as Min–Max.

To this end, we have used two of the testbeds presented in the previous chapter, namely SIGNET and Mobile testbed. These are representative of somehow typical indoor scenarios, a large room, cluttered with desks, chairs, computers, hubs, cables and people engaged in their daily work life and an empty corridor. We have then collected hundreds of RSSI measurements for each sensor and gathered all the data into two data sets, one for each environment. The data sets have been used to characterize the radio channel, through estimation of the parameters of the radio propagation models and of the random term as explained in Chapter 2.2.1. The localization algorithms have been compared over a number of different scenarios which have been extracted from the set of collected measures.

3.2.1 Localization algorithms

Out of the large variety of localization algorithms, we considered for our experiments four classical schemes, namely Multilateration, Min–Max, Maximum Likelihood-based and RO-CRSSI, which are described on the following.

Multilateration

We consider an indoor square area A containing M beacons. Let $\mathbf{b}_i = [X_i, Y_i]$ denote the geographical coordinates of the i-th beacon, for $i = 1, 2, \ldots, M$, and $\mathbf{B} = [\mathbf{b}_1; \mathbf{b}_2; \ldots; \mathbf{b}_M]$ be the $M \times 2$ matrix of all beacon coordinates. The stray (i.e., non localized) node can occupy any position $\mathbf{s} = [x, y]$ in the area, according to a given bi-dimensional probability density function $f(\mathbf{s}) : A \mapsto \mathbb{R}$, for which it clearly holds $\iint_A f(\mathbf{s}) \partial \mathbf{s} = 1$.

Multilateration [32] [33] is a simple localization algorithm, range-based, decentralized and based on geometry. Every node in position $\mathbf{s}_i = [x_i, y_i]$ collects a set of received power measurements $\mathbf{P} = \{\rho_1, \rho_2, ...\rho_M\}$ from M beacons whose positions are known. From this

information it is possible to determine the relative distances $\hat{\mathbf{d}} = \{\hat{d}_1, \hat{d}_2, ...\hat{d}_n\}$ from every beacon using (2.10)

Be $\tilde{\mathbf{s}}_i = [\tilde{x}_i, \tilde{y}_i]$ a position candidate for node i, we can define the estimation error as $e_{i,j} = |\tilde{\mathbf{s}}_i - \mathbf{b}_j| - \hat{d}_{i,j}$, i.e. the difference between the distance node–beacon inferred using the candidate position estimation $\tilde{\mathbf{s}}_i$ and the distance estimated usign the ranging technique.

The final estimate node's position $\hat{\mathbf{s}}_i = [\hat{x}_i, \hat{y}_i]$ is chosen in order to minimize the mean squared error of the $e_{i,j}$ over the entire set of the beacons. Given N beacons, the algorithm hence yields

$$\hat{\mathbf{s}}_i = [\hat{x}_i, \hat{y}_i] = \arg\min_{[x,y]} \frac{\sum_{j=0}^{N} e_{i,j}^2}{N} \tag{3.1}$$

A drawback of this technique is that every measure has the same weight, though the uncertainty in the distance estimate is larger for lower power values.

Min - Max

Another very simple algorithm is *Min-Max* [32] [33]. As Multilateration, it is a range-based, decentralized and geometric algorithm. As before, every node collects a set of received power measurements $\mathbf{P} = \{\rho_1, \rho_2, ...\rho_n\}$ from M beacons and, hence, estimates its distance from every beacon using (2.10).

At this point every node draws M squares, centered on each beacon and with side equal to twice the estimated node–beacon distance. The final position estimation $\hat{\mathbf{s}} = [\hat{x}, \hat{y}]$ is the center of the intersection of every squares. This method creates some problems when the number of beacons is large, because it is highly probable that squares have empty intersection. In this case, instead of the intersection it is possible to determine a quadrilateral in this way:

- the up side is the uppest side among the low sides of all squares
- the low side is the lowest side among the up sides of all squares
- the right side is the rightest side among the left sides of all squares
- the left side is the leftest side among the right sides of all squares

Then, the estimation is chosen as the center of this quadrilateral. This extension enables the use of this algorithm in any case. Furthermore, it is required a very light computational effort.

Once again, measures are not associated to any weight function.

Maximum Likelihood (ML)

Maximum Likelihood [27] [34] is a range-based localization scheme based on a probabilistic approach. It can be centralized or decentralized. In our testbed, we implement the distributed method.

The general formulation of a ML estimation is

$$\hat{\theta} = \arg\max_{\theta} P[\theta|\xi] \tag{3.2}$$

where θ is the vector of parameters to estimate and $\xi = [\xi_1, \cdots, \xi_M]$ is the observations vector. A more convenient formulation of the ML criterion is obtained by applying Bayes rule, which gives

$$\hat{\theta} = \arg\max_{\theta} \frac{P[\xi|\theta]P[\theta]}{P[\xi]} \tag{3.3}$$

Clearly, the estimate depends on the observations vector ξ. Therefore, we can define the mean estimation error $\epsilon(\theta)$ as

$$\epsilon(\theta) = \mathrm{E}_{\xi}\Big[||\theta - \hat{\theta}||^2\Big]$$

where $\mathrm{E}_{\xi}[\cdot]$ denotes the statistical expectation operator with respect to ξ. In our case, θ consists in the stray node's coordinates $\mathbf{s} = (x, y)$, whereas the ξ_i values correspond to the RSSI or ToA signals received from each of the M beacons leading to this expression

$$\hat{\mathbf{s}} = \arg\max_{[x,y]} \frac{P[\mathbf{P} = \rho|\tilde{x} = x, \tilde{y} = y, \mathbf{B}]P[\tilde{x} = x, \tilde{y} = y]}{P[\mathbf{P} = \rho]} \tag{3.4}$$

After some easy algebra, we find the following expressions of the ML estimator for the RSSI and ToA case:

$$\begin{aligned}
\hat{\mathbf{s}}^{rssi} &= \arg\min_{\tilde{\mathbf{s}} \in A} \sum_{j=1}^{M} \left[\log\left(\frac{\hat{d}_j^{rssi}}{d_j(\tilde{\mathbf{s}})} \right) \right]^2 \\
&= \arg\min_{\tilde{\mathbf{s}} \in A} \sum_{j=1}^{M} \log(d_j(\tilde{\mathbf{s}})) \log\left(\frac{d_j(\tilde{\mathbf{s}})}{(\hat{d}_j^{rssi})^2} \right) \qquad (3.5) \\
\hat{\mathbf{s}}^{toa} &= \arg\min_{\tilde{\mathbf{s}} \in A} \sum_{j=1}^{M} \left[\hat{d}_j^{toa} - d_j(\tilde{\mathbf{s}}) \right]^2 \\
&= \arg\min_{\tilde{\mathbf{s}} \in A} \sum_{j=1}^{M} d_j(\tilde{\mathbf{s}}) \left(d_j(\tilde{\mathbf{s}}) - 2\hat{d}_j^{toa} \right) \qquad (3.6)
\end{aligned}$$

where $d_j(\tilde{\mathbf{s}}) = ||\tilde{\mathbf{s}} - \mathbf{b}_j||$ is the Euclidian distance between the j-th beacon and the hypothetical position $\tilde{\mathbf{s}}$ of the stray node and $\hat{d}_j^{rssi|toa}$ is the estimated distance from the j-th beacon using the RSSI or ToA method.

ROCRSSI

ROCRSSI, originally proposed in [35], is a range-free algorithm, that only relys upon the assumption that P_r is a decreasing function of the distance between transmitter and receiver. Every beacon node collects RSSI measurements from all the other beacons located in its coverage area. A stray node S, located in $\mathbf{s}$, performs the following steps to estimate its own location:

1. Get all the $\rho_{\mathbf{b}_i,\mathbf{b}_j}$ measurements collected by beacon $\mathbf{b}_i$ from beacon $\mathbf{b}_j$ for each pair (i, j).

2. Collect $\rho_{\mathbf{b}_i,S}$ measurements from each beacon node $\mathbf{b}_i$ located within the coverage area of S

3. For every beacon node $\mathbf{b}_i$ in range:

   (a) find all the beacon nodes O_j for which holds $\rho_{\mathbf{b}_i,O_j} < \rho_{\mathbf{b}_i,S}$ and from them select the beacon O for which $\rho_{\mathbf{b}_i,O} = max_j(\rho_{\mathbf{b}_i,O_j})$

   (b) find all the beacon nodes I_j for which holds $\rho_{\mathbf{b}_i,I_j} > \rho_{\mathbf{b}_i,S}$ and from them select the beacon I for which $\rho_{\mathbf{b}_i,I} = min_j(\rho_{\mathbf{b}_i,I_j})$

   (c) state that S is located inside the ring centered on $\mathbf{b}_i$ having inner radius $d_{\mathbf{b}_i,I}$ and outer radius $d_{\mathbf{b}_i,O}$. If O_j is empty, no ring is traced, if I_j is empty the inner radius is zero.

4. Find the region where the largest number of rings intersect

5. Estimate the position of S in the center of gravity of this region.

3.2.2 Results

In this section we show the experimental results obtained running the different localization algorithms in the two different scenarios early discussed. The aim of this analysis is to explore how the localization precision is affected by the number of beacon nodes and by the specific algorithm used.

Fig. 3.1 shows a map of the variance of the ML localization error in an area corresponding to the SIGNET testbed, for a fixed deployment of six beacons. The color of each cell corresponds to the average value of the localization error observed in that cell. Notice that, unlike in the other graphs reported here that always refer to empirical data, the results shown in Fig. 3.1 have been obtained by simulating different realizations of the radio channel, according to the estimated parameters. Observing the results reported in the figure, we cannot conclude much about the dependence of the localization error on the beacon positions. As expected, localization error tends to reduce in proximity of the beacons, although this is not always true. Conversely, the cells close to the room corners experience larger localization errors, probably due to the worse multipath effects. However, the behavior of the localization error in indoor environments remains rather unpredictable.

In Fig. 3.2 and Fig. 3.3 we report the average localization error, defined as the distance between actual and estimated position, obtained using the different algorithms in the first and the second testbed, respectively. It is worth to notice that in Mobile testbed we have moved a single beacon node in several different positions (each corresponding to a *virtual*

beacon), so that we cannot apply the ROCRSSI algorithm that requires RSSI measurements between different pairs of beacons. The results are represented with their statistical confidence intervals, which grow when the number of beacons increases. In fact, the statistical confidence of the results depends on the number of different scenarios that could be extracted from our testbed deployments. Therefore, the larger the number of beacon nodes, the lower the number of different topologies that could be extracted and, consequently, the lower the statistical confidence of the results.

At first, we can notice that the performance obtained by the different algorithms in the two environments is qualitatively comparable, with ML outperforming all the other algorithms when the number of beacon is equal to 7 or more. When the number of beacons is less than 7, however, the other algorithms yield comparable or even better performance. It should be noted that the ML algorithm has a much higher computational cost than the other schemes here considered, and it requires a good characterization of the radio channel, including the standard deviation σ_Ψ of the shadowing term. The Min-Max scheme, on the contrary, is much less demanding in terms of computational cost and does not require the characterization of the shadowing term. Unfortunately, the Min-Max performance does not improve significantly by increasing the number of anchor nodes. In fact, the scheme presents a performance floor that apparently depends on the considered environment.

The performance of the ROCRSSI and multilateration algorithms are comparable, and generally worse than the other schemes. However, ROCRSSI offers the advantage of not requiring any radio channel characterization. In general, for the same number of beacon nodes, the performance obtained in Mobile testbed is better than that achieved in SIGNET Testbed. This indirectly confirms that the presence of furniture and moving people exacerbates the RSSI–based localization problems.

Finally, Fig. 3.4 shows the cumulative distribution function (CDF) of the localization error obtained in SIGNET Testbed, using ML and Min-Max schemes, with 5, 15, 25 and 35 beacons. It is interesting to observe that the CDF curves of ML grow more smoothly than those of Min-Max. In other terms, the localization errors obtained by adopting the ML algorithm span from a few centimeters to a few meters, whereas the Min-Max algorithm tends to concentrate the localization errors in a smaller region. This is probably due to the well known tendency of Min-Max to shift position estimates towards the center of the network [33], thus limiting the estimation error to half of the room side.

3.2.3 Discussion

In this study, we analyzed the behavior of various localization algorithms, namely ML, Min-Max, Multilateration and ROCRSSI, using real data from two different scenarios. We show that the ML algorithm yields better performance than the others when the number of anchor

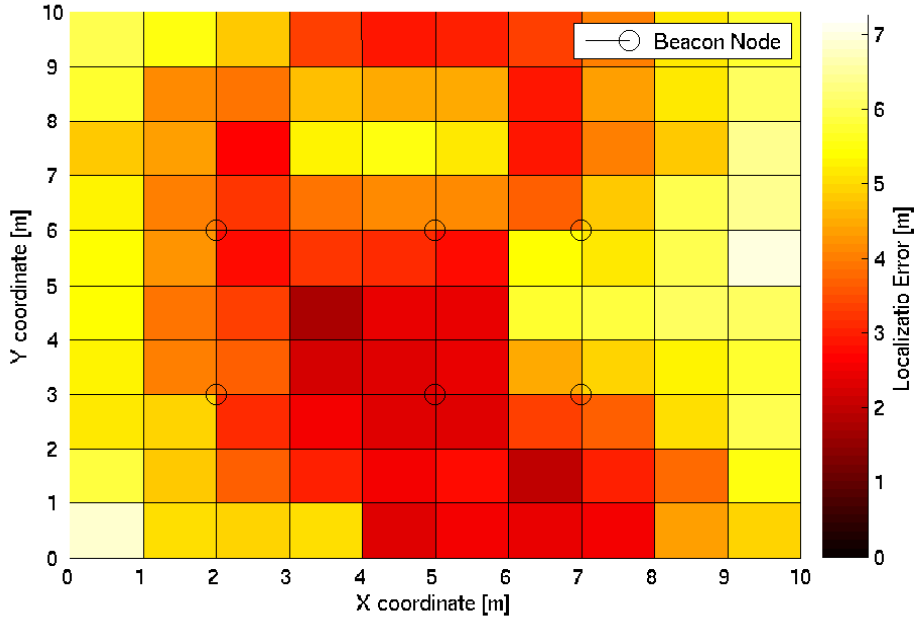


Figure 3.1: Localization error map in SIGNET testbed (simulated)

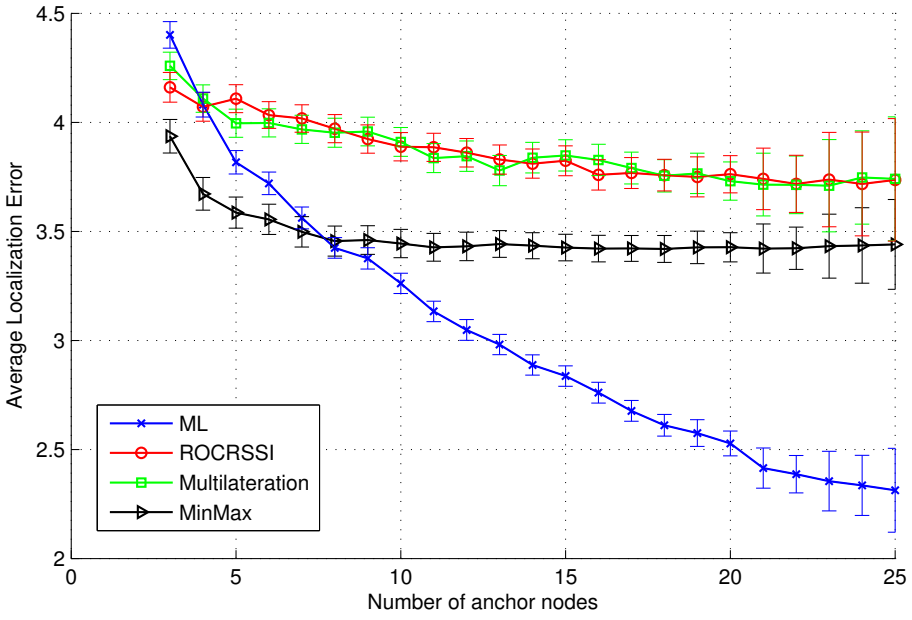


Figure 3.2: Localization error in SIGNET Testbed

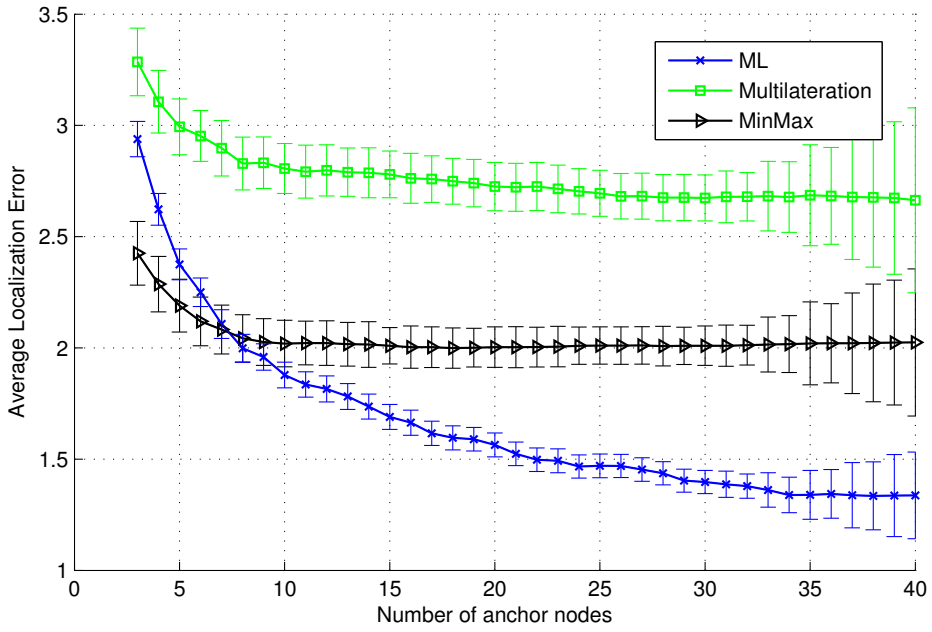


Figure 3.3: Localization error in Mobile Testbed

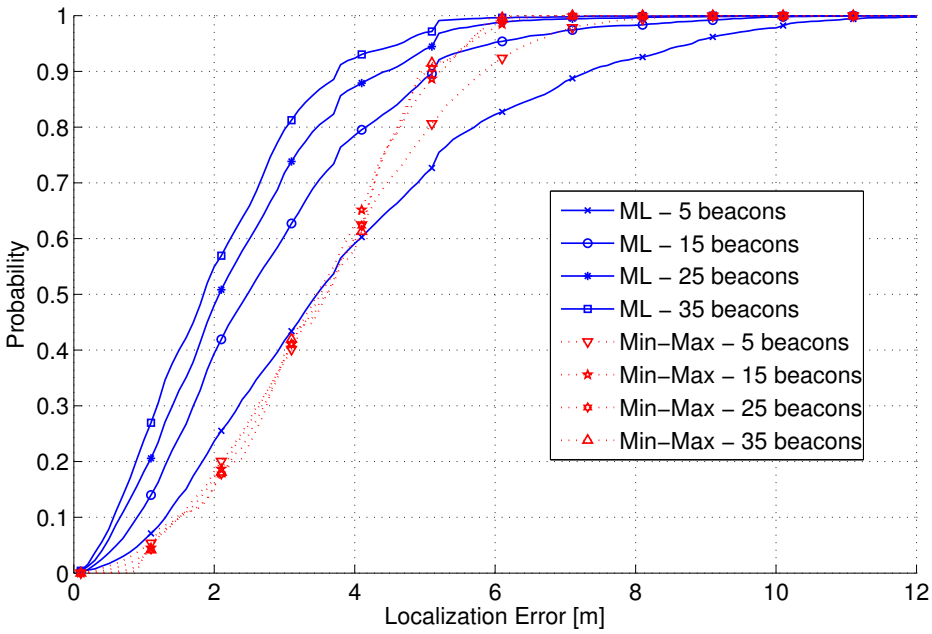


Figure 3.4: Cumulative distribution function of the localization error on SIGNET Testbed

nodes is relatively high. The main reason is that ML is the only algorithm that weights the RSSI measurements received from the anchor nodes according with the reliability of such a reading, which decreases along with the received power. However, the performance obtained in our testbeds are still rather unsatisfactory, with localization errors comparable to the inter–node distance. Multilateration is much simpler than ML, though it achieves even worse performance. The same considerations hold true for ROCRSSI, which however offers the advantage of being independent of the channel parameters estimates. Finally, Min-Max seems to offer a good compromise, having a very low computational cost and offering better results than Multilateration and ROCRSSI. However, the relatively good performance achieved by the Min-Max algorithm is mainly due to its tendency to localize the stray nodes in the center of the area, thus limiting the distance error to half of the maximum distance between the two farthest locations in the considered area.

In conclusion, RSSI-based localization in indoor environments presents severe limitations. An accurate radio channel model might alleviate these problems, though the presence of moving people or obstacles would further exacerbate the situation. Therefore, in most common indoor scenarios, RSSI-based localization schemes do not appear suitable to provide accurate localization (with errors limited to few centimeters) by leveraging on only a limited number of beacons deployed in the area.

3.3 Optimal beacon positioning

The localization problem has been primarily tackled from the perspective of the stray node, thus investigating and designing schemes and hardware that may reduce the positioning error from the signals received from beacons. In this section, instead, we turn our attention to the impact of the beacon positioning on the performance of the localization algorithms. In particular, we investigate how the mean square error of the position estimation is affected by the number and location of the beacons in the area.

The study is performed by considering different node's density distribution in the area. The problem can be reduced to the minimization of a multi-dimensional non-linear function that requires numerical methods, and whose complexity grows exponentially with the number of beacons, thus making exhaustive solution unfeasible. Therefore, we also propose an empirical method that achieves quasi-optimal performance in all the considered cases by keeping a linear complexity.

As before, we consider an indoor square area A containing M beacons. Let $\mathbf{b}_i = [x_i, y_i]$ denote the geographical coordinates of the i-th beacon, for $i = 1, 2, \ldots, M$, and $\mathbf{B} = [\mathbf{b}_1; \mathbf{b}_2; \ldots; \mathbf{b}_M]$ be the $M \times 2$ matrix of all beacon coordinates. The stray node can occupy any position $\mathbf{s} = [x, y]$ in the area, according to a given bi-dimensional probability density function $f(\mathbf{s}) : A \mapsto \mathbb{R}$, for which it clearly holds $\iint_A f(\mathbf{s}) \partial \mathbf{s} = 1$.

The localization performance depends on the observations used by the ML estimator. In our study, we consider both RSSI-based and ToA-based localization approaches that, besides being widely studied, also correspond to two different ranging-error laws as explained above.

According to (3.5) and (3.6), the mean square estimation errors is $\epsilon(\mathbf{s}) = \mathrm{E}\left[||\mathbf{s} - \hat{\mathbf{s}}||^2\right]$, where $\hat{\mathbf{s}}$ is the estimate position of the node. Averaging over the position of the stray node, we finally get the *overall mean square error* (OMSE) of the position estimation

$$\varepsilon(\mathbf{B}) = \iint_{(x,y)\in A} \epsilon(x,y) f(x,y) dx dy, \tag{3.7}$$

where we have made explicit in the notation the dependency of the OMSE on the beacon position matrix $\mathbf{B}$. Our aim is to determine the matrix $\mathbf{B}_{opt}$ that minimizes (3.7), i.e.,

$$\mathbf{B}_{opt} = \arg\min_{\mathbf{B}\in A} \varepsilon(\mathbf{B}). \tag{3.8}$$

Unfortunately, this problem does not allow for analytical solution, so that we need to resort to numerical methods.

3.3.1 Beacons positioning schemes

In order to understand the impact of the beacons position on the OMSE index, we considered three positioning strategies, namely *random*, *classic*, and *(quasi)-optimal*.

The *random* strategy simply consists in placing the beacons at random in the area. We include this limit case in our study in order to have a reference value that permits to better appreciate the performance gain obtained from an accurate planning of the beacons' position.

In the *classic* strategy, beacons' placement is driven by common-sense criterion, according to which beacons shall be preferably placed along the perimeter, specifically in the corners and the center of the edges delimiting the area, or equally spaced within the area. In order to cover most of the arrangements considered in the literature, we consider the following two strategies that we name Classic 1, and Classic 2. In Classic 1, beacons occupy the center of the regions obtained by an equal division of the area in M parts. In Classic 2, instead, beacons are placed first at the corners, then in the center of each edge. Furthermore, if there are an odd number of beacons, the last one is always placed in the center of the area. Fig. 3.5 shows an example of the two classic beacon arrangements with 7 beacons.

Finally, the *(quasi)-optimal* strategy consists in positioning the beacons according to the matrix $\mathbf{B}_{opt}$ that solves (3.8). The prefix "quasi" is motivated by the fact that, as previously mentioned, the solution requires numerical methods that do not guarantee optimality. In fact, in order to make tractable the problem, we split the area A in a regular grid of $N \times N$ cells and we limit our search to the points in the center of such cells. An exhaustive search of the solution space of (3.8) requires to consider $\binom{N^2}{M}$ possible ways in which the M beacons can be placed in the grid of $N \times N$ cells, with a complexity of the order of $O(N^{2M})$. Therefore,

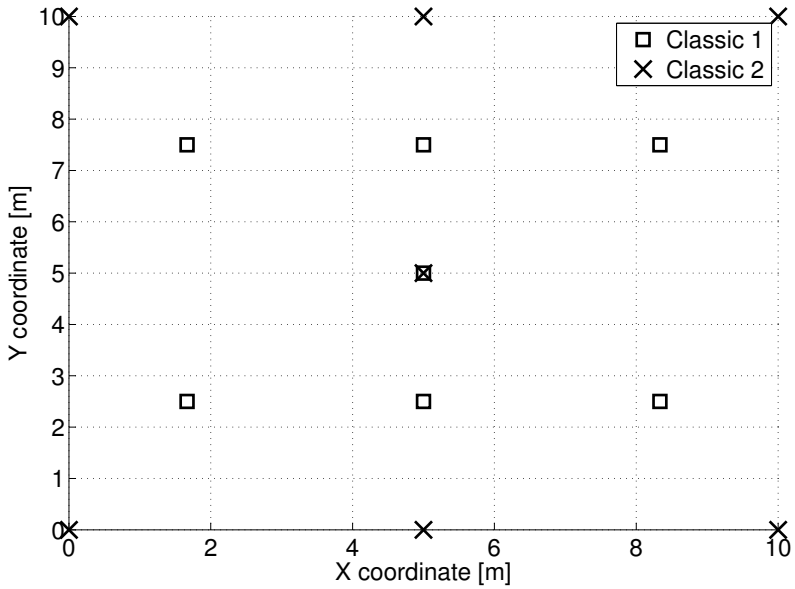


Figure 3.5: Classic 1 and 2 schemes for 7 and 8 beacons

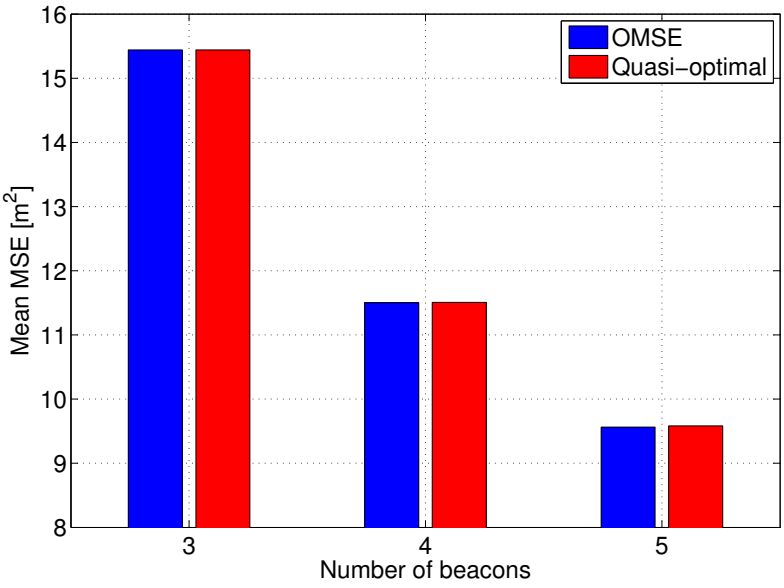


Figure 3.6: OMSE for the exhaustive search (left) and IBP (right) with uniform distribution of the stray node in the area.

the exhaustive search is practically feasible only for 3 or 4 beacons, whereas it becomes impractical for higher number of beacons. To overcome this limitation, we propose a heuristic scheme, called *Iterative Beacon Positioning* (IBP), that attempts to reduce the OMSE in successive rounds of refinement, following a sort of gradient descendant principle. Basically, IBP moves one beacon at a time, in cyclic order, minimizing at each step the OMSE obtained by considering the positions of the other beacons as static. A round consists in a cycle of beacons shift. The algorithm is, then, iterated for a finite number C of rounds, where C can be either fixed or decided at run time according to the OMSE obtained at each round. A step requires to evaluate (3.7) for each of the N^2 possible positions where the beacon under process can be moved. Since a round consists of M steps, the complexity of IBP algorithm is $O(MN^2)$, linear in M. There is no need to say that such an algorithm is not guaranteed to converge to the optimal value, since it visits only a small subset of the solution space. As a matter of fact, by applying IBP in the few cases where (3.8) can be solved by exhaustive search, the final positions returned by IBP generally do not correspond to the optimal ones (even though they are rather close). Nonetheless, the OMSE value obtained in the two cases is practically the same, as shown in Fig. 3.6 for 3, 4 and 5 beacons. These results suggest that the function (3.7) has several local minima whose value is rather close to the absolute minimum, that is to say that many different dispositions of the beacons in the area lead to similar and quasi-optimal OMSE values. Therefore, although IBP does not provide any guarantee of optimality, it seems to be successful in finding solutions that are fairly close to the optimal one. Therefore, whenever an exhaustive search of the solution space is not feasible, we will settle for labeling IBP results as "quasi-optimal".

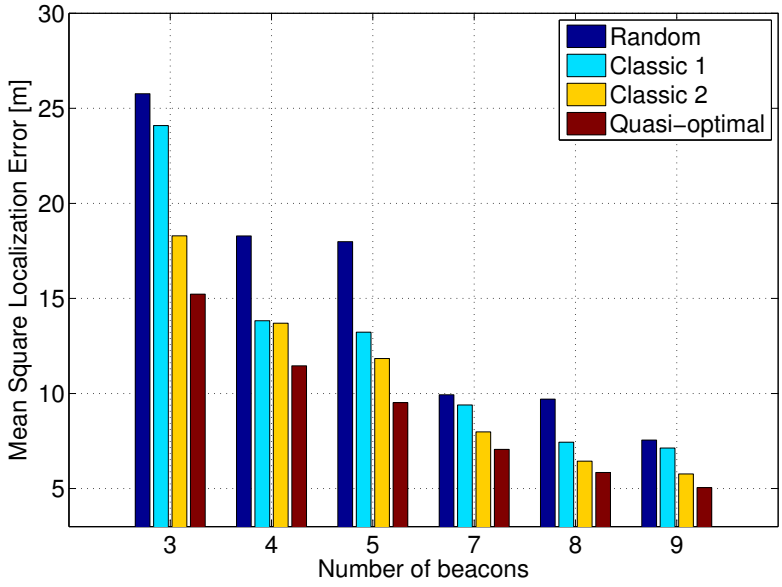
3.3.2 Results

In this section, we study the OMSE obtained by placing the beacons according to the *random*, *classic*, and *(quasi)-optimal* strategies defined in the previous section. Results have been obtained by considering an area A of 10×10 square meters, divided into $N^2 = 100$ cells of 1×1 square meters each. The considered channel model is the one described in (2.7) with the following parameters: $K = -30.5$ dBm, $\gamma = 2$, $d_0 = 0.1$ m, $\sigma = 6$ dB; whereas for the ToA model (2.14) we assumed $v_p = 3 \cdot 10^8$ m/s, $\mu_T = 10.9 \cdot 10^{-9}$ s, $\sigma_T = 3.1 \cdot 10^{-9}$ s.

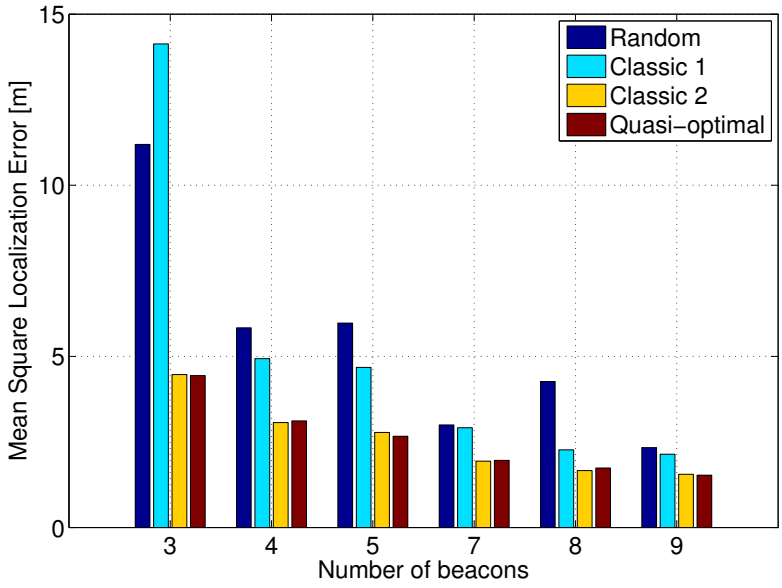
In the following, we first present the results obtained with uniform distribution of the stray node in the area and, successively, we discuss the scenario with non-uniform distribution. In both cases, we consider RSSI-based and ToA-based localization.

Uniform distribution

In this scenario, we assume that a stray node can occupy any position in the area, with equal probability.



(a) RSSI-localization



(b) ToA-localization

Figure 3.7: OMSE for different positioning schemes with RSSI-based localization (a) and ToA-based localization (b)

Fig. 3.7 reports the OMSE obtained with the four different beacons placement schemes, both in case of RSSI-based localization (upper graph) and in case of ToA-based localization (lower graph). The positions determined by the IBP algorithm are shown in Fig. 3.8. First

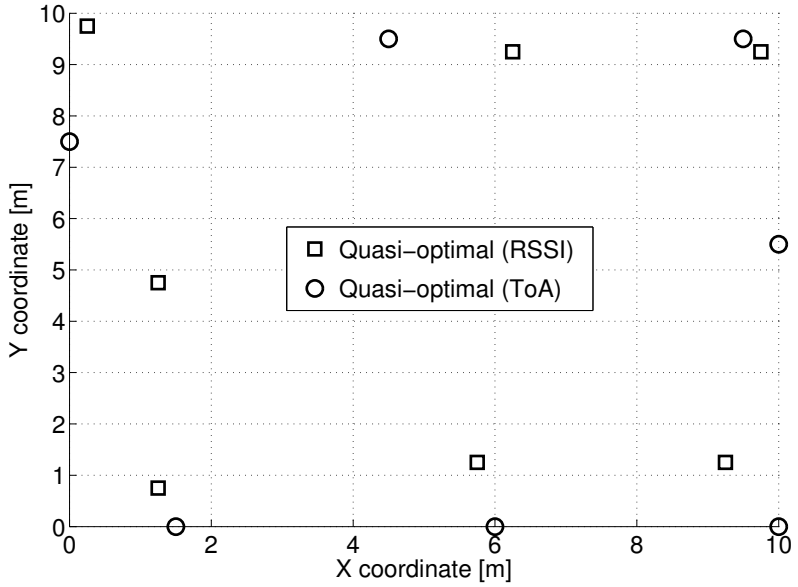


Figure 3.8: Quasi-optimal beacons position with uniform node distribution, for RSSI and ToA localization

of all, we can immediately note that, as expected, the random strategy yields much worse performance than the others. This result confirms that smart beacons positioning can greatly improve the localization performance, in particular with a small number of beacons. Moreover, we note that classic strategies perform in a similar manner, with Classic 1 achieving slightly better performance than Classic 2. The most interesting result, however, concerns the performance obtained with the quasi-optimal strategy. In fact, we can notice that we have a gain of approximately 10%–20% over the classic schemes with RSSI-based localization, whereas there is no gain with ToA-based localization. Observing Fig. 3.8, we can also note that quasi-optimal strategy tends to place the beacons at the border of the region in the ToA case, and slightly away from the perimeter and toward the center of the area in the RSSI case.

The reason of such a behavior lies on the dependance of ranging accuracy on the distance. In fact, according to (2.13), the RSSI–based localization error is generally much lower in regions close to the beacons. Moving the beacons towards the center of the area, therefore, improves the localization in the middle of the area while worsening the performance at the periphery. Therefore, the optimal position balances the two effects to minimize the OMSE.

Conversely, the accuracy of ToA-based localization is independent of the distance between beacon and stray node, as expressed by (2.15). In this case, the localization error is primarily affected by the geometrical setup of the beacons with respect to the stray node, rather than by their distance to it. Intuitively, the more sparse the beacons around the stray node,

the more selective the ML function around the real position of the stray node. Conversely, moving the beacons toward the center of the area, the points in the periphery would have all the beacons along a single direction, so that the area in which the ML function is close to its minimum would enlarge. Placing the beacons at the border of the area guarantees that each point within the area receives signals from different directions, thus improving the localization performance.

Non–uniform distribution

Here we consider a scenario in which the stray node is not uniformly distributed in the area. This might be the case, for example, of an open field, where the nodes density might depend on the conformation of the terrain. In this situation, the ML function is calculated by taking into consideration the *a priori* knowledge on the distribution $P[\theta]$ of the unknown parameters vector θ. As a case study, we consider the node's density shown in Fig. 3.9, obtained as convex combination of three bi-dimensional gaussian distributions centered in $(9.5, 9)$, $(0.5, 0.25)$ and $(5.3, 2)$ and with variances $\sigma_G^2 = [8.5\ 2.25\ 0.75]$, respectively.

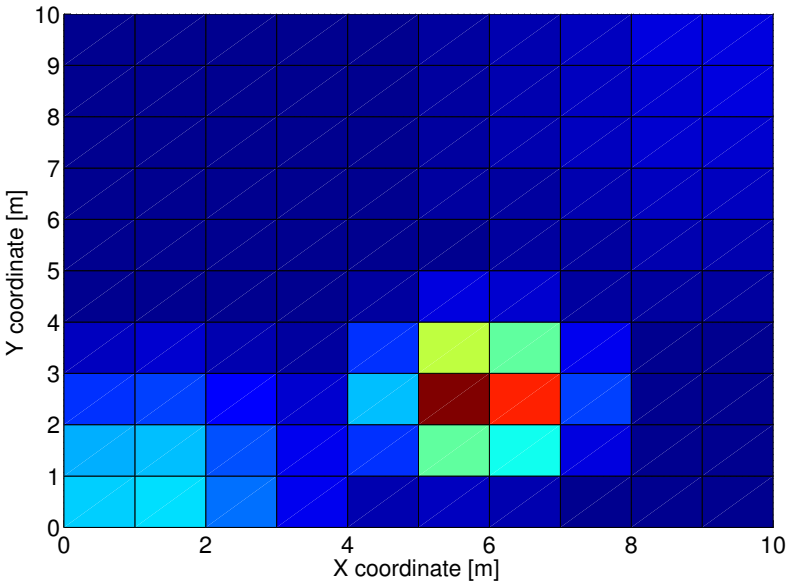
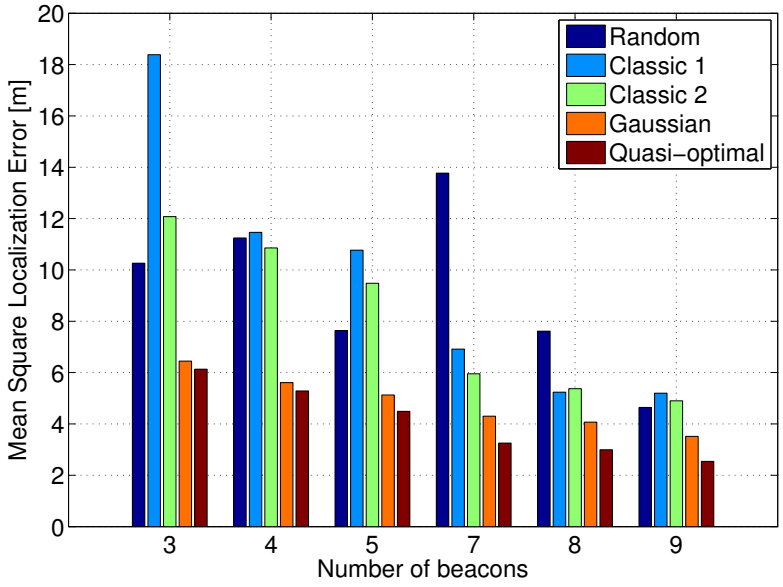


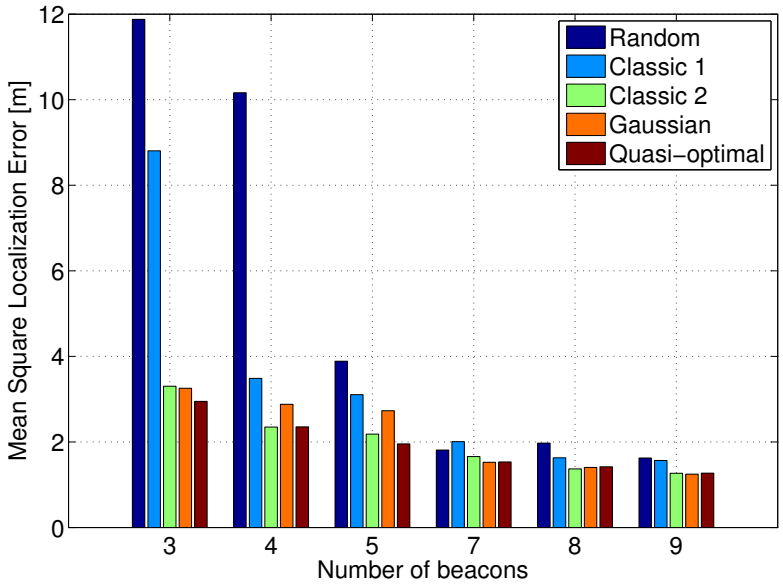
Figure 3.9: Gaussian mixture distribution

Besides the Classic 1 and Classic 2 schemes, in this scenario we also consider a "Gaussian" scheme that places the beacons in the centers of the gaussian bells that form the mixed distribution and, then, at the corners of the area.

Fig. 3.10 reports the OMSE obtained with the five different schemes for beacons positioning, for RSSI-based and ToA-based localization. The positions determined by the IBP algorithm in the two cases are reported in Fig. 3.11. We can notice that, when some regions



(a) RSSI



(b) ToA

Figure 3.10: Comparison between a classical position and the heuristic algorithm different runs with RSSI for 3, 4, 5, 7, 8 and 9 beacons

have higher probability of hosting stray nodes, the OMSE given by random, Classic 1 and Classic 2 strategies is significantly worse than the quasi-optimal one. The improvement is approximately at 45%–50%. Also, the gaussian scheme, which takes into consideration the

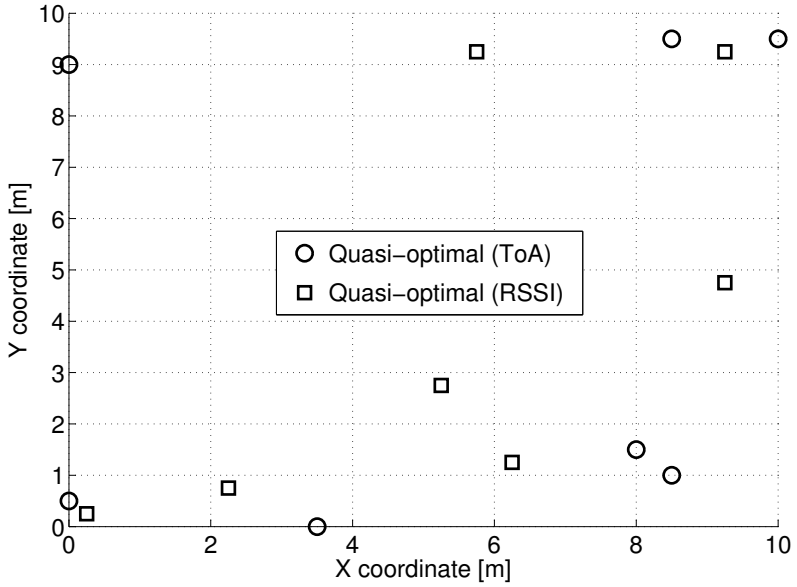


Figure 3.11: Quasi-optimal beacons position with non-uniform node distribution, for RSSI and ToA localization

non-uniform distribution of the stray node in placing the beacons, achieves good performance, with OMSE rather close to the quasi-optimal one. In this case our proposed approach for beacons positioning achieves an accuracy improvement from 5 to 25%. Looking at Fig. 3.11, we can see that, for RSSI-localization, the quasi-optimal strategy places some beacons at the corner of the area, in order to provide some basic localization quality over the entire region, and others closer to the regions with higher node's probability. According to what observed in the uniform scenario, the quasi-optimal beacons arrangement with ToA-localization still consists in placing the beacons at the perimeter of the area.

3.3.3 Discussion

The study has revealed that beacons positioning may significantly affect the localization error in a given area, both with RSSI-based localization and ToA-based localization. More specifically, an efficient beacon positioning strategy for RSSI-based localization consists to find a trade-off between placing the beacons to the periphery of the area in order to homogeneously cover the area and placing the beacons inside the area to minimize the distances among nodes and beacons beacuse of the strong dependance between distance and ranging accuracy in RSSI-based technique. Conversely, the most effective strategy with ToA-based localization consists only in placing the beacons on the periphery of the area. In fact, being the ToA ranging error independent on the distance between beacon and stray node, the localization performance is affected by the shape of the polygon formed by the beacons around the stray

node and not by the distance of the beacons.

3.4 SLAM and localization with mobile nodes

The cooperation among a WSN and an AMR can lead to many advantages in terms of localization. The robot usually has less hardware constraints in terms of computation and localization devices. In addition, the mobility of the robot enables to have communications from different locations in the network: if the AMR has some knowledge about its location, it can act as multiple virtual beacons, providing a much better localization accuracy to sensor nodes as described in Sec. 3.2. We exploited these features, studying the SLAM algorithm. This technique is necessary in scenarios where there is no "a priori" knowledge of the environment. The robot can estimate the distance to nearby nodes of the Wireless Sensor Network by measuring the Received Signal Strength Indicator (RSSI) of the received radio messages. We adopted an Extended Kalman Filter in the SLAM algorithm to integrate RSSI measurements from different nodes over time, while the robot moves inside the environment.

Results are obtained by using the generic path loss model (2.7) for the trasmission channel. Moreover, no internode communication is required in the WSN. This grants energy saving and enables the use of the proposed system also in fully disconnected networks.

The experiments have been performed using the Mobile testbed as presented in Sect. 2.3.2.

Filtering noisy measurements

The distance between the robot and each node is estimated through (2.9). All the collected measurements are reported in Fig. 3.12. The path loss model parameters K and η have been estimated from the collected RSSI measures according to a mean square error criterion. The reference distance is set as $d_0 = 10$ cm. The deterministic path loss model parameters are $P_t + K = -30.5$ dB, $\eta = 1.5$. The shadowing term Ψ_i, as expected, can be fairly well approximated by a Normal random variable with mean $\mu_\Psi \simeq -0.0348 \pm 0.0860\,\text{dB}$ and standard deviation $\sigma_\Psi \simeq 6.339 \pm 0.0614\,\text{dB}$, where the range corresponds to the 95% confidence interval. The resulting model is plotted as a blue solid line in Fig. 3.12. As seen in the figure, RSSI samples may be way too high or low with respect to the expected value for a given distance. To limit these effects, we pre-filtered the RSSI measurements taking into account the odometry information. If the robot moved a short distance between two consecutive measures, the second RSSI measure cannot change too much from the first measure.

Fig. 3.13 plots the improvements on the RSSI estimation obtained thanks to the pre-filtering of the actual RSSI measurements in a typical experiment. The blue line is the measured RSSI, the red line is the RSSI obtained after the pre-filter, the magenta line is the expected RSSI considering the channel model (2.8) and the actual distance between the robot

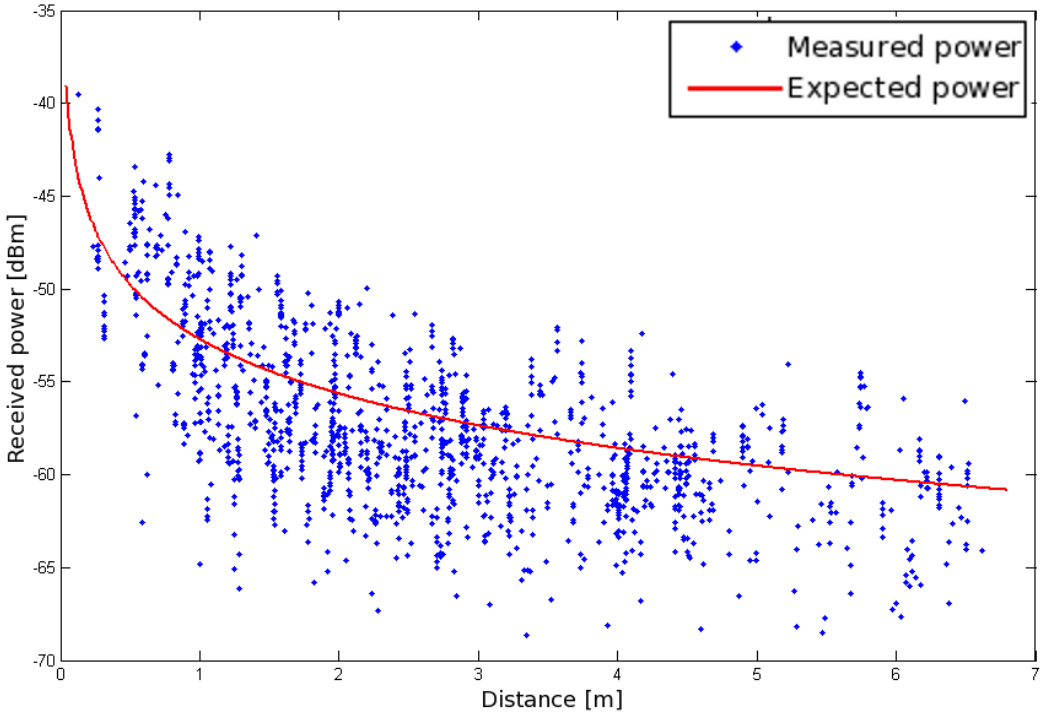


Figure 3.12: RSSI measures of the robot from a WSN node as a function of the distance robot-node. The blue solid line is the path loss model (2.8)

and the transmitting node. One can see that pre-filtering reduces the variations with respect to the expected value. In particular, several peaks and valleys are evened out. Extensive experiments proved the effectiveness of this simple filter.

3.4.1 SLAM Algorithms

The SLAM algorithm we considered is based on an Extended Kalman Filter (EKF) and it is very similar to the one presented in [38], from which we borrow the formalism. The robot state (pose and heading) at time k is $q_k = [x_k, y_k, \theta_k]^T$. The motion model of the system is:

$$q_{k+1} = \begin{bmatrix} x_k + \Delta D_k \cos(\theta_k) \\ y_k + \Delta D_k \sin(\theta_k) \\ \theta_k + \Delta\theta_k \end{bmatrix} + \nu_k = f(\hat{q}_k, u_k) + \nu_k, \tag{3.9}$$

where ν_k is a noise vector, ΔD_k is the distance traveled, and $\Delta\theta_k$ is the heading change both measured by the robot's odometers. The system state matrix $A(k)$ is given by the Jacobian:

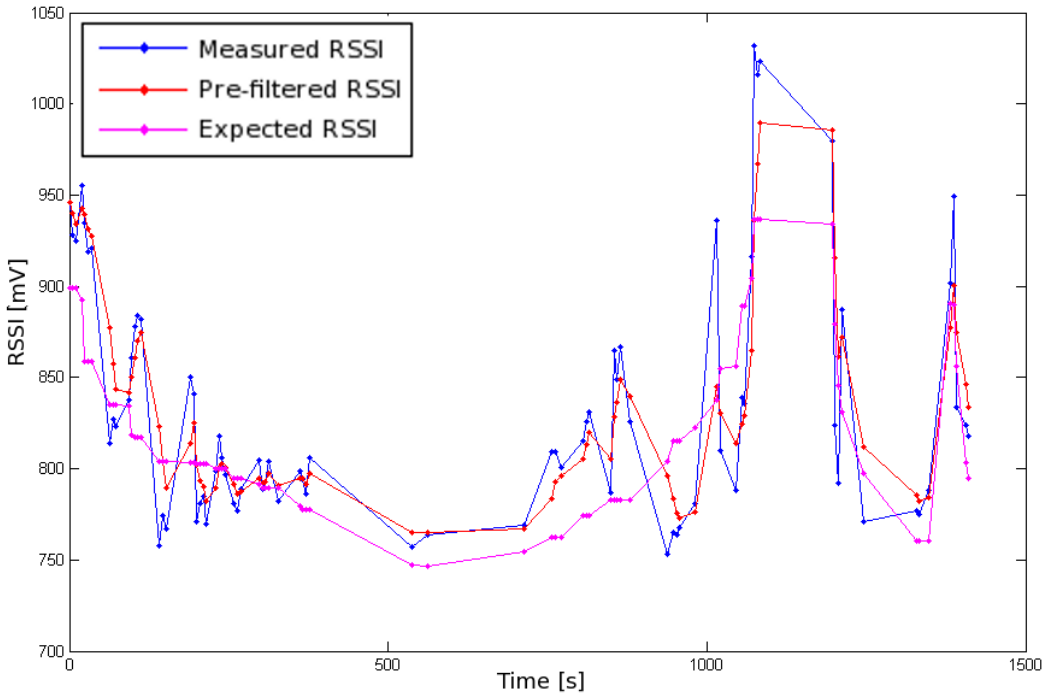


Figure 3.13: The effect of RSSI pre-filtering in a typical experiment.

$$A\left(k+1\right)=\left.\frac{\partial f}{\partial q_k}\right|_{q=\hat{q}_k}=\begin{bmatrix}1 & 0 & -\Delta D_k\cos\left(\theta_k\right)\\ 0 & 1 & \Delta D_k\sin\left(\theta_k\right)\\ 0 & 0 & 1\end{bmatrix} \tag{3.10}$$

The input matrix $B(k)$ is given by the Jacobian:

$$B\left(k+1\right)=\left.\frac{\partial f}{\partial u}\right|_{q=\hat{q}_k}=\begin{bmatrix}\cos\left(\theta_k\right) & 0\\ \sin\left(\theta k\right) & 0\\ 0 & 1\end{bmatrix} \tag{3.11}$$

When a new input from the odometry u_k is available, the robot state is updated with motion equation (3.9). The standard equations of the EKF updates the covariance matrices. When a new range measure between the robot in the state q_k and the node in the position (x_n, y_n) is available, it can be expressed as:

$$h\left(q_1\,,\,\left[x_n\,,\,y_n\right]^T\right)=\sqrt{\left(x_k-x_n\right)^2+\left(y_k-y_n\right)^2} \tag{3.12}$$

By linearizing with the Jacobian, we can write:

$$H\left(k\right)=\left.\frac{\partial h}{\partial q}\right|_{q=\hat{q}_{k+1}}=\begin{bmatrix}\frac{x_k-x_n}{\sqrt{\left(x_k-x_n\right)^2+\left(y_k-y_n\right)^2}} & \frac{y_k-y_n}{\sqrt{\left(x_k-x_n\right)^2+\left(y_k-y_n\right)^2}} & 0\end{bmatrix} \tag{3.13}$$

Tha above formulas are for the robot localization only, but in the case of SLAM also the wireless node must be localized so the new vector of state of the system is:

$$q_k = \begin{bmatrix} x_k \\ y_k \\ \theta_k \\ x_{n,1} \\ y_{n,1} \\ \vdots \\ x_{n,i} \\ y_{n,i} \end{bmatrix} \tag{3.14}$$

Also the matrices $A(k)$ and $B(k)$ must be modified and they become:

$$A\left(k+1\right) = \left.\frac{\partial f}{\partial q_k}\right|_{q=\hat{q}_k} = \begin{bmatrix} 1 & 0 & -\Delta D_k \cos\left(\theta_k\right) & 0 & & 0 \\ 0 & 1 & \Delta D_k \sin\left(\theta_k\right) & 0 & & 0 \\ 0 & 0 & 1 & 0 & & 0 \\ 0 & 0 & 0 & 1 & & 0 \\ & & & & \ddots & 0 \\ 0 & 0 & 0 & 0 & & 1 \end{bmatrix} \tag{3.15}$$

$$B\left(k+1\right) = \left.\frac{\partial f}{\partial u_k}\right|_{q=\hat{q}_k} = \begin{bmatrix} \cos\left(\theta_k\right) & 0 \\ \sin\left(\theta_k\right) & 0 \\ 0 & 1 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix} \tag{3.16}$$

Now the time update does not change with respect to what is written in (3.9), but for the measurement update the Jacobian $H(k)$ must be calculated considering that also the wireless node position has to be updated. Note that only terms non zero in $H(k)$ are the ones relative to the robot position and those relative to the node corresponding to the current measure.

Initial estimation of node position

At the start-up, the EKF needs to be initialized with a *first guess* of the node position. This information was obtained thorugh trilateration that provides an estimation $\hat{\mathbf{s}}$ of an object position, by knowing its distance from M objects (provided they do not lie on a line). This is done by looking for the intersection of the circumferences centered in the reference objects and with radius equal to the distances from the target object. We used least-square method to find the intersection point and therefore the estimated position. Once the first estimate

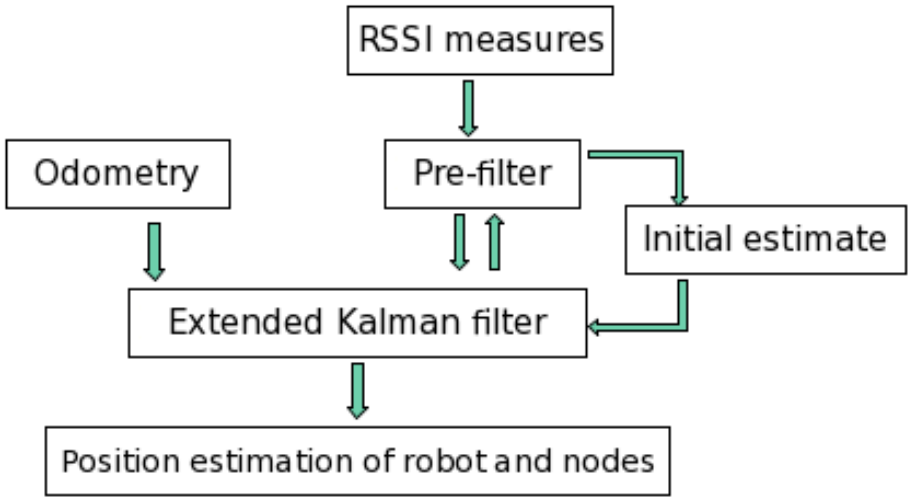


Figure 3.14: The data flow in the implemented SLAM algorithm

on the coordinates of the node is obtained, the position can be inserted in the matrices of the EKF. In particular, we expand the covariance matrix by inserting the current robot covariance summed to the mean covariance of the measures used in the trilateration.

3.4.2 Experiments

In this section we present two kind of experiments of SLAM, one assuming a prior positioning knowledge of the nodes locations; the other assuming a rough guess of the real nodes position is available. This second experiment refers to a scenario in which a human operator (or a loosely localized robot, like the helicopter reported in [39]), drops sensor nodes along a path and roughly records their positions.

The experiments were performed in the Mobile scenario (see Sect. 2.3.2), shown in Fig. 3.15. The SLAM experiment results are graphically reported in Fig. 3.16 and Fig. 3.17.

3.4.3 Discussion

From our experiments, we showed that the proposed approach, that works online updating the system state every time a new measure is received, can achieve a mean error on the position of the WSN between 0.5 m and 1.0 m.

A further result of our work is that our approach is strongly relying on the initial guess on the node estimation. In Fig. 3.18 are reported the average final errors obtained by the SLAM algorithm for robot localization and nodes' positions. With respect to Fig. 3.18, `XTE` and `ATE` are respectively the *Cross-Track* mean Error and the *Along-Track* mean Error; `Cartesian`

Figure 3.15: The robot and the WSN deployed in the Mobile testbed.

is the mean error on the Cartesian distance between the estimated robot position and the ground-truth; `Nodes` is the mean error on the Cartesian distance between the estimated and the actual nodes' positions. The series `1,2,3` in Fig. 3.18 are those referring to the experiments with the triangulation algorithm to initialize in the SLAM the position of the newly encountered nodes along the path. The series `7,8,9` in Fig. 3.18 are those referring to the experiments in which a first guess of nodes positions within 2 meter from the actual one is available. In the latter case the residual error is almost half of the one achieved in the first case. Therefore, a more robust system to estimate the first guess on the nodes position is needed.

One should also note that in Fig. 3.18 the standard deviation associated to the residual error on the nodes' positions is very large (the lines over the bars of the histogram). This is because the initial guess of the node position is strongly variable, ranging from 1 m to 4 m.

The differences within the two series of Fig. 3.18 are due to the way in which multiple measures received while the AMR is on at the same position are managed by the SLAM system. In series `1,7` when the robot is steady, only the first three measures in chronological order of reception are inserted in the SLAM algorithm. In series `2,8` when the robot is steady, only the three measures with highest RSSI are inserted in the SLAM algorithm (i.e.,

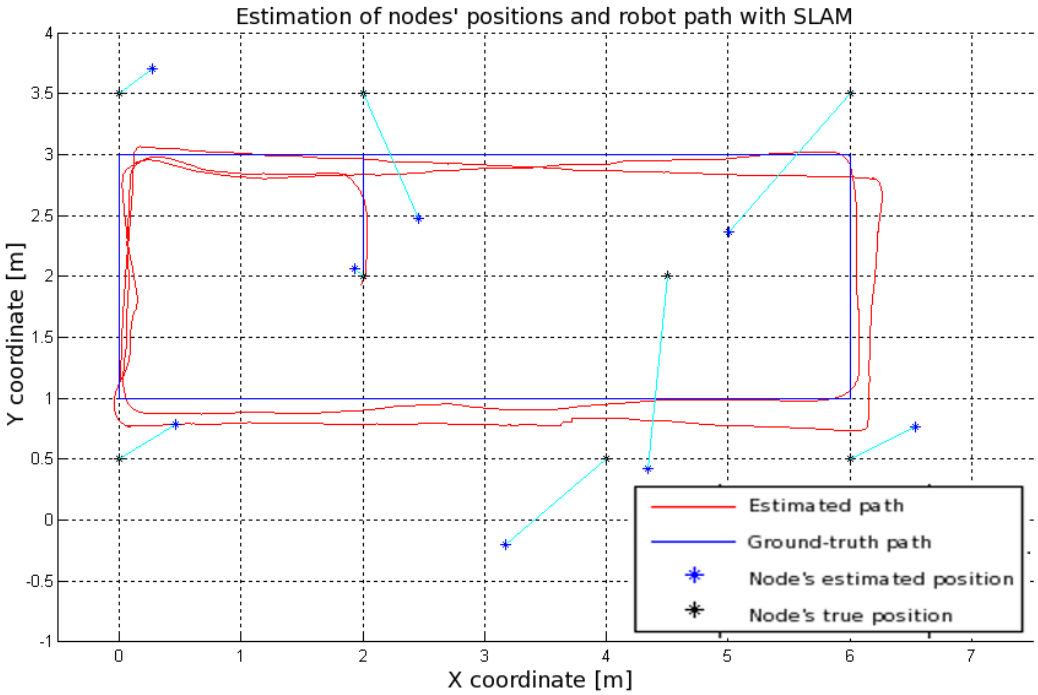


Figure 3.16: SLAM result starting from no "a priori" positioning knowledge.

the SLAM state is updated only when the robot moves again). In series `3,9` a fully off-line approach is tested where only the 40 measures with the highest RSSI for each node are inserted in the SLAM algorithm.

In conclusion, localization results are in line with the best previous works in which a robot is used as a mobile WSN node. Better improvements (almost a factor 4) have been obtained with respect to the localization algorithms when a WSN node exploits only static nodes to estimate its position. These results are very good considering that they were obtained in an indoor environment, without any calibration of the receiving/transimitting systems on the nodes, without a specific model of the actual communication channel of the environment in which the experiments were performed.

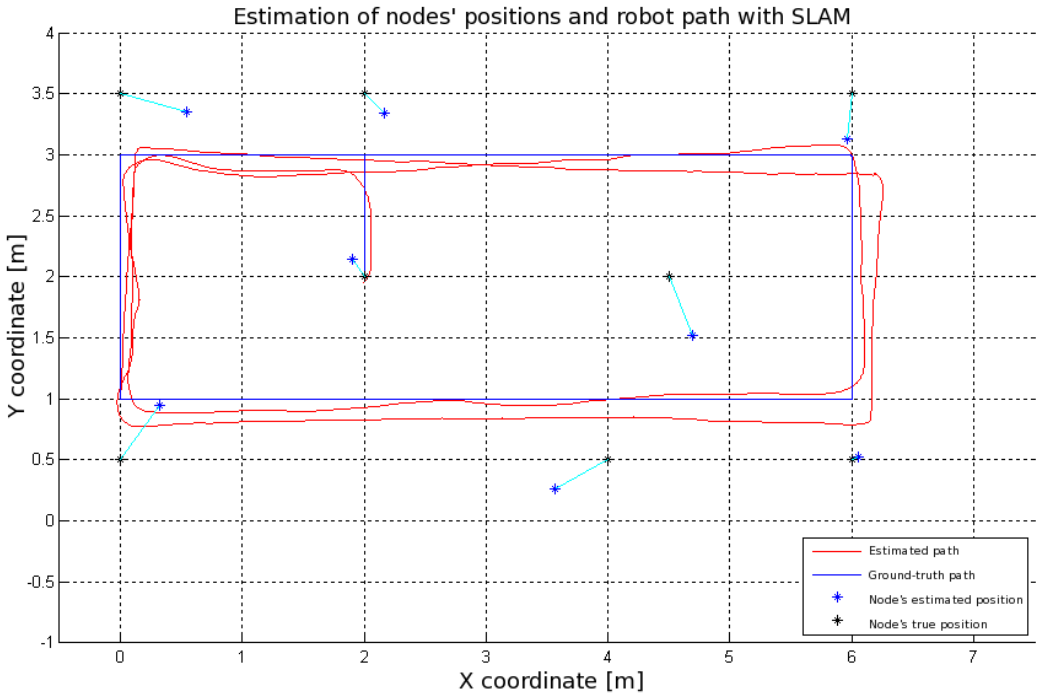


Figure 3.17: SLAM result starting from a rough initialization of the nodes' position.

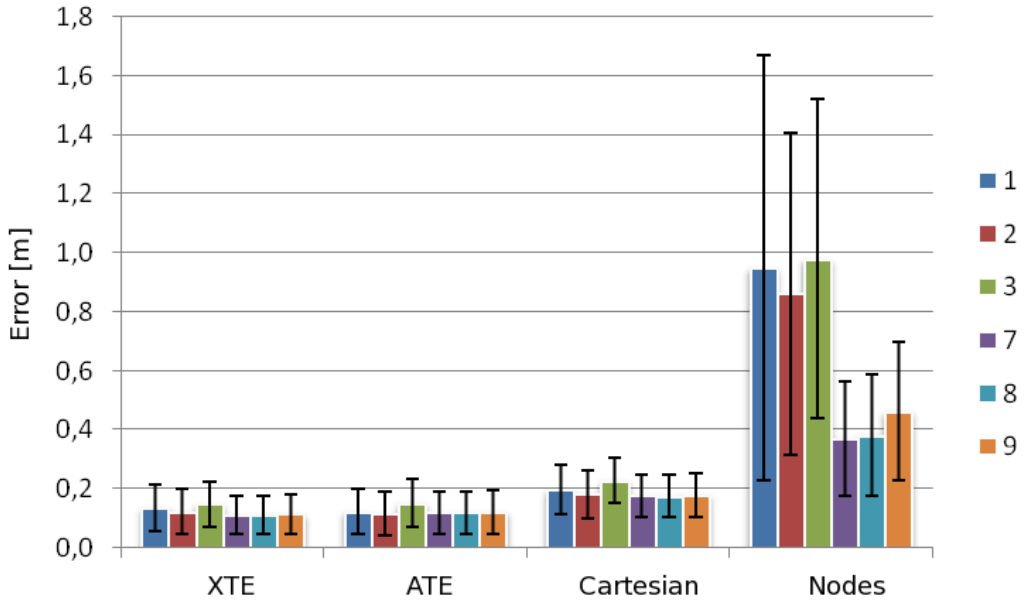


Figure 3.18: Comparison between average residual error on the robot and nodes' positions obtained by the SLAM algorithm when starting from no prior knowledge (1, 2, and 3 in the legend), and from an initial gross localization (7, 8, and 9 in the legend).

Chapter 4

Opportunistic localization

In this chapter we tackle the problem of node localization from a different perspective: rather than searching for yet another signal-processing technique or system architecture explicitly designed to provide localization services, we propose to spill out this service from the opportunistic interactions that may occur among heterogeneous wireless nodes.

The opportunistic paradigm relies on the cooperative interaction among nodes that, however, occurs on an occasional basis only. Therefore we can talk of *opportunistic interaction* to define the information exchange among nodes. A typical example of an opportunistic interaction scenario is the seamless data exchange among portable devices carried by people in public areas, such as malls, theaters, hotel lounges and so on [40]. In this case, people move in a rather uncoordinated manner, each person following his personal purpose, so that data exchange among devices carried by people can occur only occasionally.

We propose this scenario observing two main results obtained in Chapter 2 and Chapter 3. First, the degradation of the ranging accuracy with the distance, when using RSSI measurements. Second, the limit of the beacon based RSSI localization unless to use a large number of beacons and to use a high computational demanding algorithm like ML. In the opportunistic scenario, we consider a native limited localization accuracy based on some infrastructured system and we try to improve the performance using closer nodes to retrieve more reliable ranging information.

We address the problem of providing accurate self-localization service in such an opportunistic scenario. We envision a number of mobile nodes with different mobility patterns and equipped with heterogeneous communication devices and self-localization hardware, e.g. Cricket [41], MEMS [42], indoor GPS [43], RSSI-based or none. Nodes are capable of performing seamless and opportunistic data exchange to attain certain goals and, in particular, to improve their positioning estimate. More specifically, we investigate how the localization accuracy of nodes can be ameliorated by opportunistically exchanging localization information with other nodes that occasionally happen to be in proximity.

This scenario offers a number of research challenges that include the definition of efficient nodes discovery and link establishment algorithms to connect multi-interface devices, the design of suitable opportunistic data exchange protocols, the devising and analysis of localization enhancing schemes based on exchanged data.

We propose two different algorithms for improving the localization accuracy of the nodes that occasionally enter in contact with each other and exchange positioning information; one is based on a Maximum Likelihood argument, and the other on a greedy, heuristic approach. The algorithms are first modeled and analyzed from a theoretical perspective. Analytical results are then compared with measures collected from Opportunistic testbed (see Sec. 2.3.4). The rather good match between theoretical and experimental results confirms the accuracy of the analytical model, which is then used to investigate the performance of the proposed schemes in a larger set of scenarios. The results confirm that the opportunistic localization paradigm may effectively improve the localization accuracy of nodes, in particular in heterogeneous networks.

ML and GOAL algorithms assume that the node has a native self–localization estimation, provided by an infrastructure or by means of other techniques.

A very interesting case can consider that a node is completely unaware of its position then it has to rely only on the other nodes information. There are several ways to compute an accurate position based on inaccurate position estimations from peers. One of the most promising approaches is based on Linear Matrix Inequalities (LMI) and was introduced by Doherty et al. [44]. It consists in establishing a set of inequalities verified by the user and peers coordinates, which form a LMI system; if the system is solvable, any of the solutions can be used as an estimation of the user position. We analyzed the problem from different points of view.

We proposed a *range-free* LMI scheme where only a proximity information is inferred from the exchanged packet. The inequalities took into account self-positioning errors as well as a constant maximum communication range.

Then, we propose a *range-based* LMI scheme in which the Received Signal Strength (RSS) from each peer is used to estimate an upper bound of the user-peer distance. Indeed, the communicating devices that we consider can read measures of the RSS: we exploit these measures to estimate the user-peer distance using the well-known pathloss exponent and lognormal shadowing propagation models, and bound the error of this distance estimation with a 99% confidence. This per-peer maximum distance bound is then used in the LMI system.

Furthermore, when many peers can communicate with the user, the amount of inequalities in the LMI system increases and the computational effort needed to solve it also increases. We propose a Hybrid approach that trades off computational effort for accuracy by selecting subsets of inequalities before solving the system.

Finally, the proposed scheme and improvements are evaluated through Matlab simulations, in order to investigate the localization accuracy when varying the number of available contact opportunities in different scenarios. The performance achievable with this approach is analyzed in different scenarios, through extensive simulation campaign. Results show that LMI-based schemes, especially the range–based one, are potentially capable of yielding very accurate localization even after a limited number of opportunistic exchange, though performance is rather sensitive to the accuracy of the other nodes' self-localization and to the randomness of the radio channel.

4.1 State of the art

A recent research trend addresses the localization problem in cooperative scenarios, which have been extensively studied, in particular, in the robotic area. A typical reference case consists in teams of mobile autonomous robots equipped with different sensors that cooperate one another and, occasionally, interact with simple sensors placed in the environment to achieve a given goal, such as node localization and tracking. Motion tracking algorithms generally leverage on Extended Kalman Filter [45] or Particle Filter [46, 47] in order to exploit the correlation among different measurements. In [48] the authors utilize Markov localization in order to self–localize nodes and then probabilistic methods to synchronize each robot's estimates when two of them have a contact. A distributed Kalman Filter is performed for collective localization in [49], avoiding a centralized data fusion, that is not so feasible in a cooperative scenario. An anchor-free approach is proposed in [50], where robots infer their position only using the information exchanged among them. Similar approaches are proposed for very specific applications as in [51] for video surveillance and in [52] for autonomous vehicles in mining. The literature on the opportunistic interaction paradigm is mainly focused on routing and forwarding scheme as in [53, 54, 55, 56, 57].

4.2 Opportunistic meeting model

In this section, we define a mathematical model of an opportunistic interaction. The model is based on a simplified scenario that, although idealized, includes some of the most interesting design parameters, such as the maximum range of an opportunistic communication, the fraction of time devoted to the opportunistic interactions, and the relative speed between the nodes. Then, the model permits to investigate the impact of these parameters on some performance indexes that are of interest for the opportunistic localization scheme, namely the probability of occurrence of an opportunistic interaction and the statistical distribution of the distance between the nodes when such an opportunistic data exchange occurs.

Assumptions

Radio propagation is described by means of a simple unit-disk model, according to which the radio transmission is always correctly received within a distance R (*opportunistic coverage range*) from the transmitter, whereas it is not received at longer distances. Although the unit-disk model is known to be oversimplified, it permits to isolate the performance analysis from the characteristics of the radio interface that, at this stage of the work, is left generic.

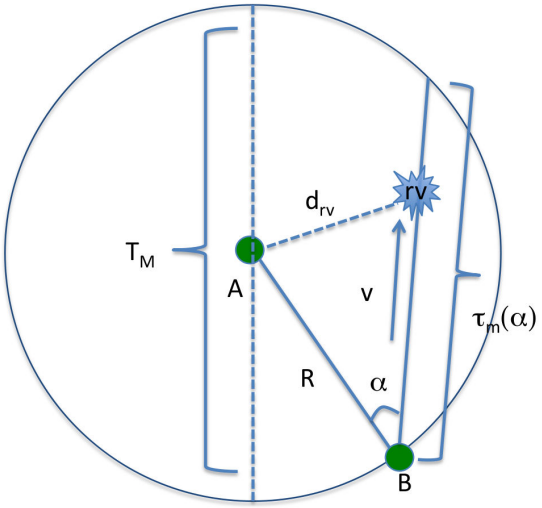


Figure 4.1: Fly-by model

A *fly-by* event occurs when two nodes enter in each other coverage range. Nodes' trajectories are straight and uniform. Hence, centering the reference system on node A, we can describe the relative trajectory of node B by means of two parameters, namely the (relative) speed v and the angle of incidence α of B's trajectory into the circle of radius R centered in A. From these parameters it is possible to infer the duration of the fly-by event. With reference to Fig. 4.1, we define the following parameters whose inter-relations can be easily obtained by basic trigonometry:

- τ: time since the beginning of the fly-by event;
- $s(\tau) = v\tau$: distance covered by B at time τ;
- $d(\tau, \alpha) = \sqrt{R^2 + s(\tau)^2 - 2Rs(\tau)\cos(\alpha)}$ Euclidean distance between A and B at time τ;
- $\tau_m(\alpha) = 2R\cos(\alpha)/v$: overall fly-by duration;
- $T_M = 2R/v = \tau_m(0)$: maximum fly-by duration.

A *rendezvous* occurs when both nodes are in *Opportunistic Scan Phase* during a fly-by, an event that we call *hit*. We denote by t_{rv} the instant when the rendezvous occurs with respect to the beginning of the fly-by. Nodes perform the Scan Phase periodically, every T seconds, in an asynchronous and independent manner, so that the offset between the scan phases of two nodes is a random variable with uniform distribution in the interval $(0, T)$. The ratio between the duration of the scan phase and the time T between two consecutive scan phases is called *duty cycle* and denoted by δ. Different choices of δ and T result in different hit probabilities and distributions of the t_{rv}.

An important role is played by the duration of the fly-by that depends on the coverage range R, the relative speed v and the incidence angle α.

4.2.1 Meeting time distribution

Let t_{rv} denote the instant at which the scan phases of the two nodes overlap for the first time, measured with respect to the beginning of the fly–by, and let $F_{t_{rv}}(\cdot)$ and $f_{t_{rv}}(\cdot)$ be the corresponding probability distribution and density functions, respectively. Due to the periodicity of the scan phases, we have $t_{rv} \in [0, T]$. When $\delta_A + \delta_B \leq 1$, there is a positive probability that the scan phases do not overlap. In this case, $F_{t_{rv}}(\tau)$ is a defective distribution with upper limit given by $F_{t_{rv}}(T) = \delta_A + \delta_B$, which corresponds to the probability of overlapping before T. After some easy algebra, the probability density function (pdf) $f_{t_{rv}}(\tau)$ turns out to be

$$f_{t_{rv}}(\tau) = \begin{cases} \delta_A \delta_B \delta(t) & \tau{=}0 \\ (\delta_A + \delta_B)/T & 0{<}\tau{\leq}T(1{-}\delta_B) \\ 2(T-\tau)/T^2 & T(1{-}\delta_B){<}\tau{\leq}T \end{cases} \tag{4.1}$$

where $\delta(t)$ is the Dirac delta function. When $\delta_A + \delta_B > 1$, the scan phases always overlap at some point in the interval $[0, T]$ and the pdf can be obtained following the same rationale explained above.

4.2.2 Rendezvous and hit probability

Given α, a rendezvous occurs when $t_{rv} \leq \tau_m(\alpha)$, so that the cumulative distribution function (cdf) $F_{t_{rv}}(t, \alpha)$ of the rendezvous time t_{rv}, conditioned on α, can be expressed as

$$F_{t_{rv}}(t, \alpha) = \begin{cases} F_{\tau_o}(t, \alpha) & \text{for} \quad 0 \leq t < \tau_m(\alpha) \\ F_{\tau_o}(\tau_m(\alpha), \alpha) & \text{for} \quad t > \tau_m(\alpha) \end{cases} \tag{4.2}$$

We note that $F_{t_{rv}}(a, \alpha)$ is a defective distribution. The upper limit of $F_{t_{rv}}(a, \alpha)$ gives to the so-called *hit probability*, that is the probability of observing a rendezvous during a fly-by,

which will be denoted by

$$P_H(\alpha) = F_{t_{rv}}(\tau_m(\alpha)) \tag{4.3}$$

Averaging over the distribution $f_\alpha(\theta)$ of α we get the expected hit probability

$$P_H = \int_{-\pi/2}^{\pi/2} P_H(\theta) f_\alpha(\theta) d\theta \tag{4.4}$$

The hit probability P_H is an important performance index, since it conveys the possibility of enabling any opportunistic algorithm in a given scenario.

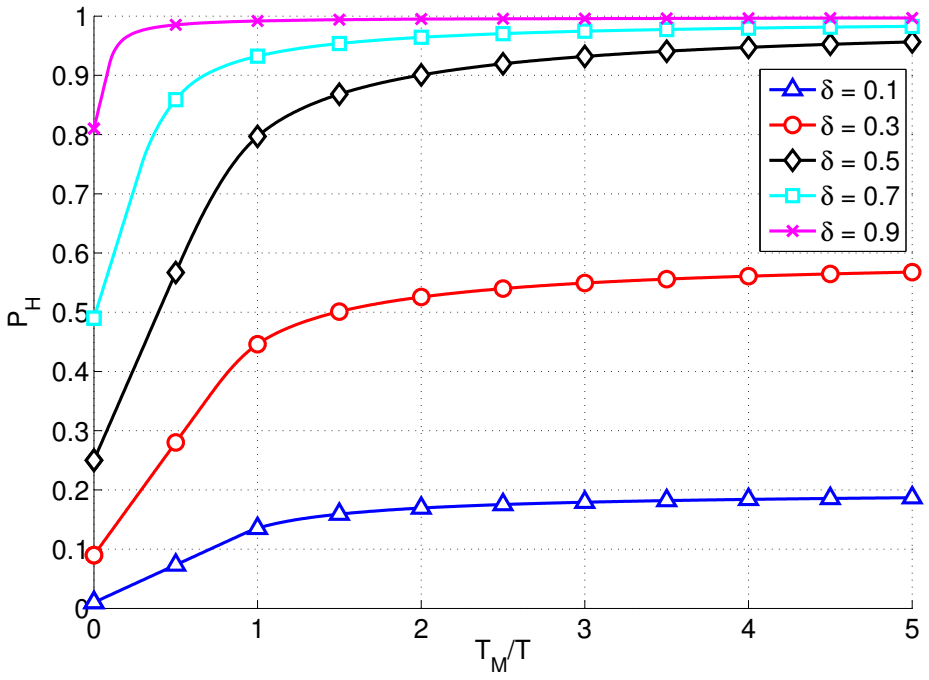


Figure 4.2: Hit probability P_H when varying the duty–cycle $\delta_A = \delta_B = \delta$

Fig. 4.2 shows P_H versus T_M/T, for different duty cycles, assuming $f_\alpha(\theta) = 1/(\pi)$ with $\theta \in [-\pi/2, \pi/2]$. The ratio T_M/T gives the *maximum* number of scan periods that the two nodes can perform during a fly-by. Note that $T_M/T > 1$ does not guarantee that the nodes always perform an entire scan phase during the fly-by, since the actual duration of each fly-by depends on α. We note that P_H grows rather rapidly till $T_M/T < 2$ after which the curves tend rather slowly to their asymptotic value $P_H = \min(1, 2\delta)$, which corresponds to the hit probability when each fly-by lasts more than T. On the light of these results, in the following we set $T_M/T = 2$.

4.2.3 Hit distance probability

The other performance index of interest for the localization scheme is the so-called "hit-distance" d, ie., the distance at which the hit occurs. Applying simple geometric arguments, we can easily realize that, given α, the nodes are at exactly at distance x during the fly-by in the two instants

$$t_{1,2}(x,\alpha) = \left(\frac{R\cos(\alpha) \mp \sqrt{x^2 - R^2\sin^2(\alpha)}}{v} \right)$$

Then, the probability that the hit-distance $d \leq x$ is equal to the probability that the rendezvous occurs in the time interval $[t_1(x,\alpha),\, t_2(x,\alpha)]$. Averaging over α, we hence get

$$F_d(x) = \int_{-\pi/2}^{\pi/2} \left[F_{t_{rv}}(t_2(x,\theta)) - F_{\tau_o}(t_1(x,\theta)) \right] f_\alpha(\theta) d\theta \tag{4.5}$$

which provides the CDF of the "hit-distance". This distribution is of great interest because the performance of range-based localization schemes closely depends on the quality of the ranging that, in turn, is a function of the real distance between the two communicating devices when using RSSI measurements.

Fig. 4.3 shows the distribution $F_d(x)$ of the hit-distance, as given by (4.5), when varying the duty–cycle δ. First of all, we observe that the curves show a discontinuity when $d/R = 1$,

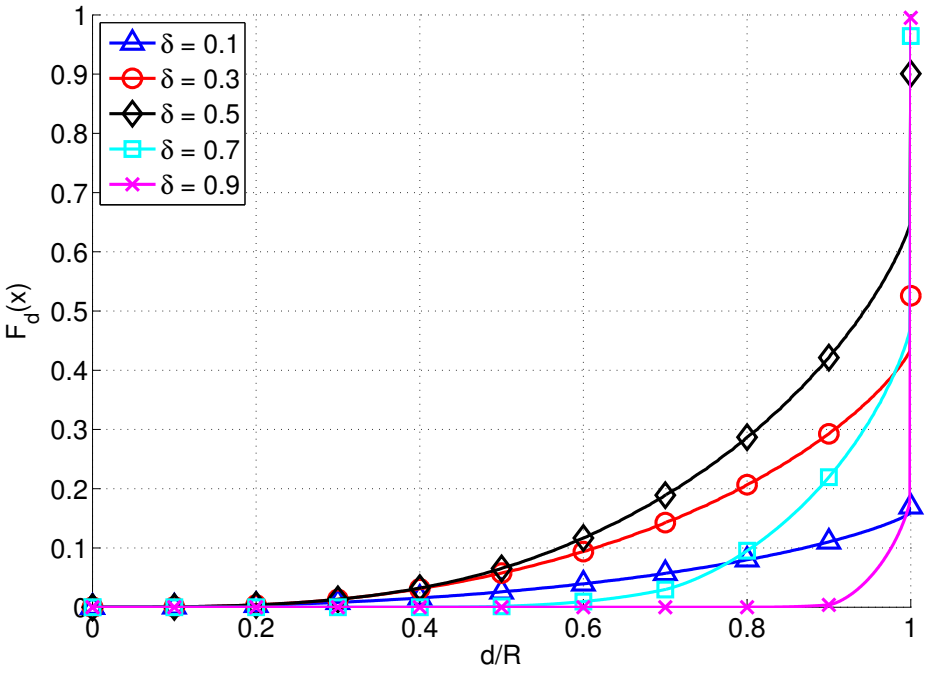


Figure 4.3: CDF of the hit-distance d when varying the duty-cycle δ

which depends on the Dirac impulse observed in (4.1). We also observe that the curves do

not reach 1, since the distribution is defective, the upper limit corresponding to the hitting probability. Recalling that RSSI-based ranging techniques usually perform better at short distance, the best setting for δ is 0.5. Smaller values of δ would reduce the hit probability, whereas larger value would increase the probability that the *rendezvous* occurs at the border of the coverage range (we consider only the first rendezvous during the fly–by).

4.3 Theoretical analysis of opportunistic algorithms

In our model, we assume that nodes have "native" self-positioning capabilities, provided by some (non opportunistic) scheme. Accordingly, we denote by $\mathbf{s}$ and $\hat{\mathbf{s}}$ the real and estimated node's position, expressed in polar coordinates, at a given time instant. For simplicity, we assume that the estimation error $\mathbf{e} = \mathbf{s} - \hat{\mathbf{s}}$ is distributed as a two–dimensional Gaussian random variable, with zero mean and variance σ^2. Accordingly, the module of the error $\epsilon = ||\mathbf{e}||$ is a Rayleigh-distributed random variable with parameter σ and probability density function (pdf)

$$f_\epsilon(x) = \frac{x \exp\left(\frac{-x^2}{2\sigma^2}\right)}{\sigma^2} \; ; \quad x \geq 0 \tag{4.6}$$

We assume that nodes can be associated to different "*positioning classes*" on the basis of the variance σ^2 that characterizes their self-localization accuracy. When two nodes, say A and B, perform an opportunistic interaction they exchange packets containing their estimated positions, $\hat{\mathbf{s}}_A$ and $\hat{\mathbf{s}}_B$, and the positioning class they belong to. The packets exchange is also used by nodes to get an estimate $\hat{d}$ of their distance, by using a ranging technique based on the RSSI using (2.9), the ToA as in (2.15) or other methods [16]. Then, each node refines the estimate of its own position to $\tilde{\mathbf{s}}$ by using an opportunistic localization algorithm. In this work we consider two different algorithms, one denoted by ML and based on a Maximum Likelihood argument and the other based on a greedy heuristic technique and called Greed Opportunistic Algorithm for Localization (GOAL).

4.3.1 ML algorithm

The ML algorithm updates the position estimate of nodes A and B to $\tilde{\mathbf{s}}_A$ and $\tilde{\mathbf{s}}_B$, respectively, by solving the following Maximum Likelihood equation

$$\tilde{\mathbf{s}}_{A,B} = \arg \max_{\mathbf{s}_A, \mathbf{s}_B} \mathrm{P}\left[\hat{d}, \hat{\mathbf{s}}_A, \hat{\mathbf{s}}_B | \mathbf{s}_A, \mathbf{s}_B\right] \tag{4.7}$$

where $\hat{\mathbf{s}}_A, \hat{\mathbf{s}}_B$ are the current position estimates of nodes A and B communicated during the opportunistic data exchange, whereas $\hat{d}$ is the distance estimate performed during such an interaction. Basically, the conditional probability on the right-hand side of (4.7) conveys the likelihood of the estimates $\hat{\mathbf{s}}_A, \hat{\mathbf{s}}_B$ and $\hat{d}$ under the hypothesis that nodes A and B are really

located in $\mathbf{s}_A$ and $\mathbf{s}_B$, respectively. The new positions are hence chosen in order to maximize such a conditional probability. Now, denoting by $f_{\hat{d}}(x)$ the conditional pdf of the ranging $\hat{d}$, given that nodes A and B are assumed to be located in $\mathbf{s}_A$ and $\mathbf{s}_B$, and recalling that the absolute positioning errors are assumed to be Rayleigh distributed, (4.7) can be written as

$$\tilde{\mathbf{s}}_{A,B} = \arg\max_{\mathbf{s}_A,\mathbf{s}_B} \left\{ f_{\epsilon_A}\left(\|\hat{\mathbf{s}}_A - \mathbf{s}_A\|\right) f_{\epsilon_B}\left(\|\hat{\mathbf{s}}_B - \mathbf{s}_B\|\right) \cdot \right. \\ \left. \cdot f_{\hat{d}}\left(\frac{\hat{d}}{\|\mathbf{s}_A - \mathbf{s}_B\|}\right) \right\} \tag{4.8}$$

where $\|\cdot\|$ denotes the Euclidean norm. Clearly, $f_{\hat{d}}(\cdot)$ depends on the specific ranging technique adopted by the nodes. In our study, we consider an RSSI-based ranging technique which provides a distance estimate given by (2.9). Hence, the ranging estimate turns out to be lognormal distributed with pdf

$$f_{\hat{d}}(x) = \frac{10\eta}{x \cdot \ln(10) \cdot \sigma_\psi \sqrt{2\pi}} e^{-\frac{(10\eta \log_{10}(\frac{x}{d_{AB}}))^2}{2\sigma_\psi^2}}, \quad x > 0 \tag{4.9}$$

and cumulative distribution function (cdf)

$$F_{\hat{d}}(x) = \Phi_{0,\sigma_\psi}\left(10\eta \cdot \log_{10}\left(\frac{x}{d_{AB}}\right)\right) \tag{4.10}$$

where $\Phi_{\mu,\sigma}(x) = \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)$ is the cdf of a normal distribution with mean μ and variance σ and σ_ψ is the standard deviation of the shadowing random term as described in (2.7).

Unfortunately, the ML equation (4.8) cannot be solved in closed form, so that we resort to Monte Carlo simulations.

4.3.2 Greedy Opportunistic Algorithm for Localization (GOAL)

The ML algorithm provides a useful benchmark for testing the performance of other opportunistic localization algorithms, being asymptotically optimal. However, the practical implementation of the algorithm, in particular in real-time scenarios, is impractical because of the very high computational cost required to solve the ML equation. For this reason, we propose a heuristic algorithm, named Greedy Opportunistic Algorithm for Localization (GOAL), which is much lighter than ML in terms of computational complexity but it offers poorer localization performance in some scenarios.

A first important difference between ML and GOAL is that, in case of opportunistic interaction between two nodes A and B, GOAL is performed only by the node with lower positioning accuracy, say node B, whereas the other node does not modify its current position estimate. Then, after data exchange, node B traces an abstract line that connects the two

current estimates of nodes' position, $\hat{\mathbf{s}}_A$ and $\hat{\mathbf{s}}_B$. GOAL updates the estimate of node B position to the point $\tilde{\mathbf{s}}_B$ that lies at distance $\hat{d}$ from $\hat{\mathbf{s}}_A$, along the conjunction line, in the direction of $\hat{\mathbf{s}}_B$. In formula, we have

$$\tilde{\mathbf{s}}_B = \hat{\mathbf{s}}_A + \hat{d}\frac{(\hat{\mathbf{s}}_B - \hat{\mathbf{s}}_A)}{||\hat{\mathbf{s}}_B - \hat{\mathbf{s}}_A||}. \tag{4.11}$$

The localization error of node B after having performed GOAL can be analytically determined. To simplify the understanding, we proceed gradually, starting with the simplest scenario and, then, relaxing progressively some simplifying assumptions. We start by assuming that both the positioning of node A and the ranging are error free. This scenario may correspond to the case in which node A is a beacon node, and ranging is performed using sophisticated techniques. The current positioning of node B, instead, is supposed to be affected by an error $\boldsymbol{\epsilon_b} = (x, y)$, where x and y are independent and identical distributed (iid) Gaussian random variables with zero mean and variance σ_B^2. Observing Fig. 4.4, which

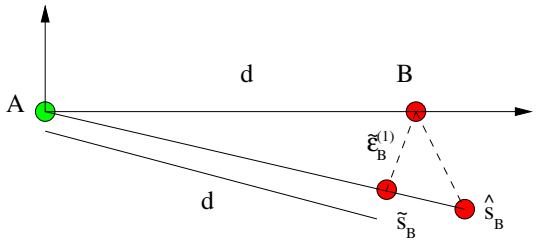


Figure 4.4: GOAL analysis: perfect node A positioning and ranging.

sketches the geometry of this simple case, it can be immediately realized that the positioning error $\tilde{\epsilon}_B$ of node B after GOAL depends only on the angle $\alpha = \widehat{BA\tilde{S}_B}$. Denoting by $f_\alpha(\cdot)$ the PDF of α and by d the true distance between A and B, the CDF $F_{\tilde{\epsilon}_B}^{(1)}(x|d) = \mathrm{P}\left[\tilde{\epsilon}_B \leq x|d\right]$ in this first case can hence be expressed as

$$F_{\tilde{\epsilon}_B}^{(1)}(x|d) = \int_0^x 2\frac{f_\alpha\left(2\arcsin\left(\frac{x}{2d}\right)\right)}{d\sqrt{1-(\frac{x}{2d})^2}}dx \tag{4.12}$$

The PDF of the angle α, in turn, can be obtained from basic trigonometry, giving

$$f_\alpha(a) = \int_{-d}^{+\infty} p_x(x)p_y((d+x)\tan(a)) \cdot (d+x)dx$$

that, after some algebra, yields

$$f_\alpha(a) = \frac{\left[1 - \mathrm{erf}\left(\frac{-d-2d\tan^2(a)}{\sqrt{2\sigma^2}\sqrt{1+\tan^2(a)}}\right)\right] e^{-\frac{d^2\tan^2(a)}{2\sigma^2(1+\tan^2(a))}}}{2\sqrt{2\pi}\sigma\sqrt{1+\tan^2(a)}}. \tag{4.13}$$

We now relax the assumption of error-free ranging by assuming that the estimated distance $\hat{d}$ differs from the real distance d for an error term ϵ_ρ. The corresponding scenario

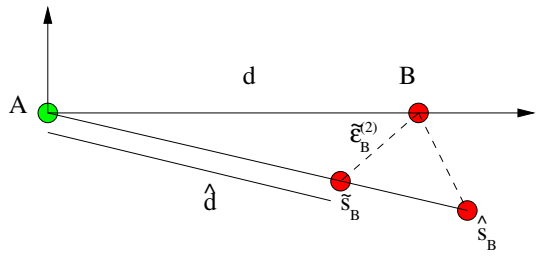


Figure 4.5: GOAL analysis: perfect node A positioning, imperfect ranging.

is represented in Fig. 4.5. The overall position error after GOAL, $\tilde{\epsilon}_B$, depends on both the error due to the angle between A and $\hat{S}_B$, equal to $\epsilon_\alpha = 2d\sin(\alpha/2)$, and the ranging error $\epsilon_\rho = \|d - \hat{d}\|$, according to the following relation

$$\tilde{\epsilon}_B^2 = \epsilon_\rho^2 + \epsilon_\alpha^2 + 2\epsilon_\rho\epsilon_\alpha \sin(\alpha/2).$$

After some cumbersome but not difficult algebra, the conditional cumulative distribution function (CDF) of $\tilde{\epsilon}$ in this second case turns out to be equal to

$$F_{\tilde{\epsilon}_B}^{(2)}(x|d) = \begin{cases} 2\int_0^{\hat{\theta}} f_\alpha(\alpha)(F_{\hat{d}}(r_2) - F_{\hat{d}}(r_1))d\alpha, & x \le d; \\ 2\int_0^{\pi} f_\alpha(\alpha)F_{\hat{d}}(r_2)d\alpha, & x > d; \end{cases} \tag{4.14}$$

where

$$r_{1,2} = d\cos(\alpha) \mp \sqrt{x^2 - d^2\sin^2(\alpha)}\,, \tag{4.15}$$

$$\hat{\theta} = \arcsin\left[\min\left(\frac{x}{d}, 1\right)\right], \tag{4.16}$$

whereas $f_\alpha(x)$ is given by (4.13) and $F_{\hat{d}}(x)$ is given by (4.10).

We can finally consider the case where even node A has imperfect knowledge of its own position, as illustrated in Fig. 4.6.

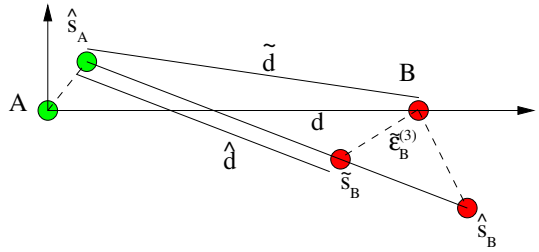


Figure 4.6: GOAL analysis: imperfect node A positioning and ranging.

The analysis of this last case can be directly derived from that of the previous case, just taking into account that the final position estimation of B now depends on the angle $\widehat{B\hat{S}_A\hat{S}_B}$

rather than on $\widehat{BA\hat{S}_B}$. It is hence sufficient to determine the statistical distribution of $\hat{d}$, whose cdf turns out to be equal to

$$F_{\hat{d}}(x) = \begin{cases} \frac{1}{\pi}\int_0^{\arcsin(\min(\frac{x}{d},1))}(F_\rho(d_2) - F_\rho(d_1))d\alpha, & x \leq d; \\ \\ \frac{1}{\pi}\int_0^{\pi} F_{\hat{d}}(d_2)d\alpha, & x > d; \end{cases} \tag{4.17}$$

with

$$d_{1,2} = d\cos(\alpha) \mp \sqrt{x^2 - d^2\sin^2(\alpha)},$$

and $F_\rho(x)$ is the CDF of a Rayleigh-distributed random variable with parameter σ_A. The cdf of the overall positioning error of B after GOAL is, hence, equal to

$$F^{(3)}_{\tilde{\epsilon}_B}(x|d) = \int_0^{+\infty} F^{(2)}_{\tilde{\epsilon}_B}(x|s)\, f_{\hat{d}}(s)ds, \tag{4.18}$$

where $F^{(2)}_{\tilde{\epsilon}_B}(x|s)$ is obtained by evaluating (4.14) in x with $d = s$, whereas $f_{\hat{d}}(s)$ is the derivative in x of $F_{\hat{d}}(x)$ given by (4.17).

As an example, we report in Fig. 4.7 the cdf of $\tilde{\epsilon}_B$ given by (4.12), (4.14) and (4.18), respectively, with $d = 2$ m, $\sigma_B = 3$ dB and $\sigma_\psi = 6$ dB (only for the second and third case) and $\sigma_A = 2$ dB (only for the third case). Using the CDF of the GOAL localization error, it is possible to infer the average localization error of GOAL when nodes are at distance d using the formula

$$E[\tilde{\epsilon}_B|d] = \int_0^{\infty}(1 - F^{(i)}_{\tilde{\epsilon}_B}(x|d))dx. \tag{4.19}$$

We can observe that GOAL, on average, improves the original localization of node B, equal to $\sqrt{\pi/2}\sigma_B$, though the gain reduces when considering non-ideal ranging and node's A positioning.

4.3.3 Validation of the theoretical model of ML and GOAL

The light computational cost required by GOAL makes it suitable for resource constrained devices, such as simple sensor nodes. To assess the performance of the algorithm in a real environment, we have implemented GOAL on the Opportunistic testbed described in Sec. 2.3.4 and run an extensive experimental campaign.

The model parameters have been determined using (2.21) from the RSSI measures gathered by the opportunistic node during the experiments and are reported in Tab. 4.1.

P_{TX}	d_0	K	η	σ_ψ
-1 dBm	10 cm	-36 dB	1.74	6 dB

Table 4.1: Experimental characterization of the radio propagation model.

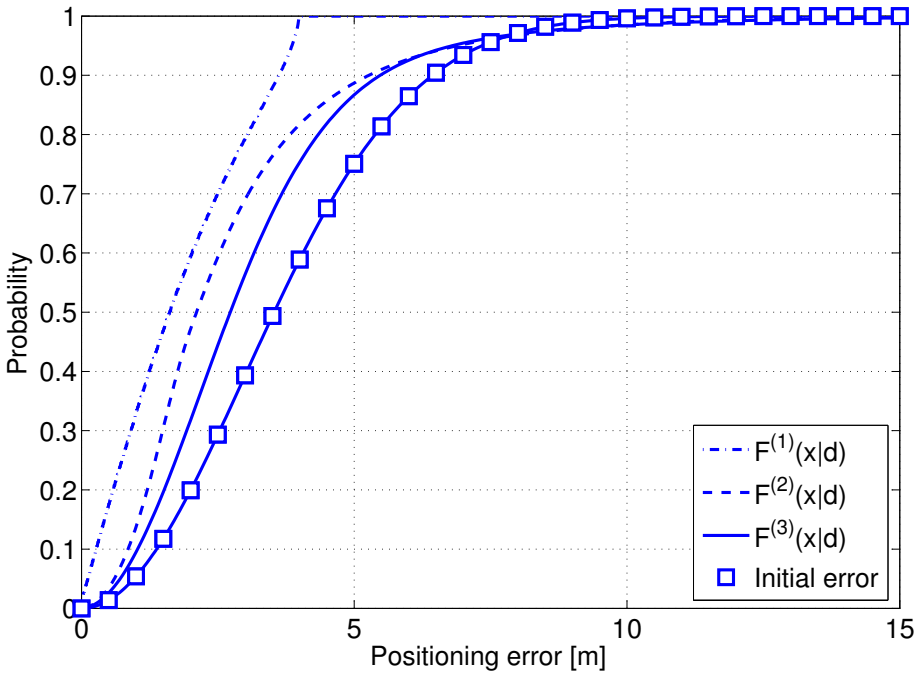


Figure 4.7: CDF of error positioning using GOAL for the three different scenarios

For each of the 120 prefixed positions along the path followed by the opportunistic node (B), it turns on its radio transceiver and runs the Min-Max localization algorithm, using the signals broadcasted by the beacon nodes to get an estimate $\hat{\mathbf{s}}_B$ of its own position. We observed that the localization error obtained in this way is approximately Rayleigh distributed (gaussian in cartesian coordinates), with overall standard deviation $\sigma_B \simeq 3$ meters. Then, node B transmits a probe message to establish an opportunistic link with the static nodes within reception range, which play the role of node A during the opportunistic interactions. Nodes reply to the probe by sending their own position estimate and the variance of their self-localization error. Note that, the position estimate $\hat{\mathbf{s}}_A$ of a static node is artificially generated by the node itself by adding an error with bidimensional gaussian distribution to its actual location. In this way, it is possible to emulate different classes of nodes by simply changing the variance σ_A^2 of the gaussian distribution. Besides the positioning data, the nodes also collect the Received Signal Strength Indicator (RSSI) for each received packet, which is used to perform ranging estimate. Data and ranging measures collected by the nodes along the path are, then, post-elaborated according to the ML and GOAL algorithms in order to determine the amelioration of the positioning error brought by the opportunistic algorithms.

The comparison between theoretical and experimental results is performed by considering

the *opportunistic gain* metric Δ_i, $i = A, B$, defined as

$$\Delta_i = \frac{\sigma_i\sqrt{\pi/2} - \tilde{\epsilon}_i}{\sigma_i\sqrt{\pi/2}}, \quad i = A, B, \tag{4.20}$$

where $\tilde{\epsilon}_i$ is the mean localization error of node i after the opportunistic localization, whereas $\sigma_i\sqrt{\pi/2}$ is the mean localization error of the same node obtained by using the native localization scheme. Therefore, Δ represents the relative gain in the localization error obtained by using the opportunistic scheme. In order to have a fair comparison, theoretical results are obtained by setting the parameters of the analytical model with the same values obtained from the experimental measures and reported in Tab. 4.1.

Fig. 4.8, Fig. 4.9 and Fig. 4.10 show the opportunistic gain Δ_B for node B obtained from the theoretical model (solid line) and the experimental measures (square marks), when varying the distance d between nodes A and B during the opportunistic interaction. Clearly, the experimental data are collected only for a finite number of positions, so that the density of marks is not homogeneous along the horizontal axis. Fig. 4.8 has been obtained by setting $\sigma_A = 0.1$ m, whereas in Fig. 4.9 we have set $\sigma_A = 2$ m and in Fig. 4.10 $\sigma_A = 3$ m. We can note the rather good agreement of theoretical and experimental results. Results for the ML algorithm, which are not reported to reduce clutter, yield similar conclusions.

Therefore, the experimental measures confirm the reliability of the theoretical model, which can then be used to further analyze the performance of the two algorithms in more complex scenarios. In particular, in the next section we complete the study by investigating the impact of the opportunistic interaction model derived in Sec. 4.2 on the performance of the opportunistic localization algorithms.

4.3.4 Opportunistic localization analysis

In this section, we investigate and compare the performance of the two opportunistic update algorithms, ML and GOAL as a function of the distance between the nodes when the opportunistic interaction takes place. Then, we extend the analysis by considering the model of opportunistic interaction developed in Sec. 4.2.

ML vs GOAL without meeting model

We consider two different cases: a heterogeneous scenario in which nodes A and B belong to different localization classes, with $\sigma_A = 0.1$ m and $\sigma_B = 3$ m, respectively; and an almost homogeneous scenario where both nodes have similar localization capability, with $\sigma_A = 2$ m and $\sigma_B = 3$ m.

Fig. 4.11 shows the results for the heterogeneous scenario (upper solid line and square markers) and the almost homogeneous scenario (lower solid line and triangular markers). We note that the opportunistic localization can effectively provide a large performance gain

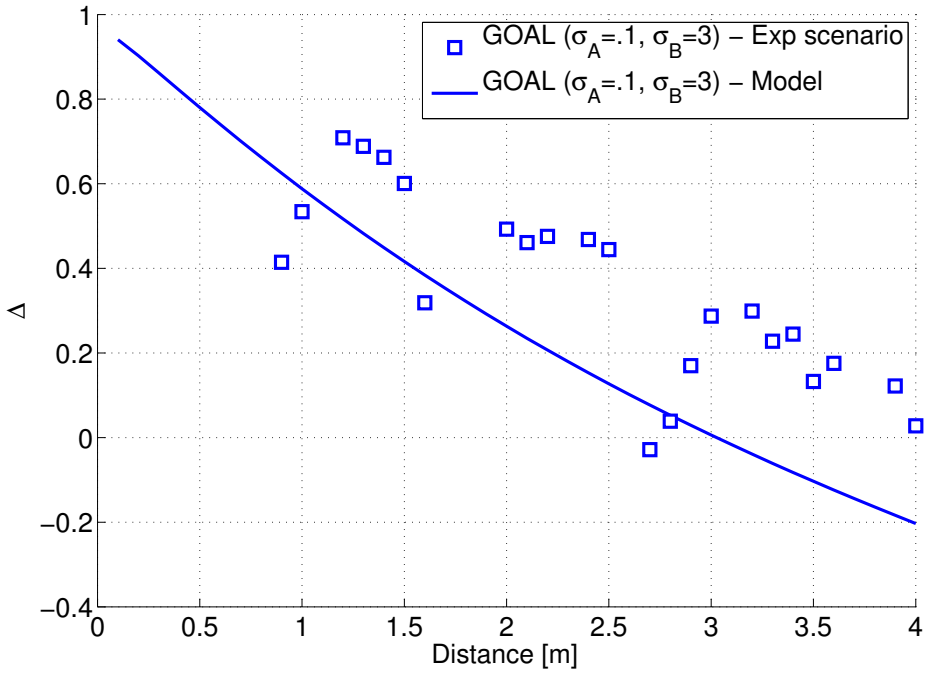


Figure 4.8: Theoretical vs experimental results for GOAL ($\sigma_A = 0.1$ m, $\sigma_B = 3$ m)

with respect to native localization system, in particular if nodes belong to different classes of accuracy and the opportunistic interaction occurs at short distances. First of all, when self–localization of node A is better than that of node B, GOAL performs slightly better than ML up to a certain distance, after which ML outperforms GOAL. The reason is that GOAL is really effective when the positioning error of A and the ranging error are both much smaller than B positioning error, whereas ML performance is penalized by the fact that the likelihood function is (negatively) affected by the rather large error in B positioning. On the contrary, when the accuracy of nodes A and B are similar, or the distance between the node is large, then ML outperforms GOAL. Moreover, it is very interesting to observe that ML, on average, does not worsen the initial accuracy of nodes' positioning, even in a very harsh scenario, whereas GOAL may even increase the native positioning errors of the nodes in certain cases.

ML vs GOAL with meeting model

In order to complete the comparison between the two algorithms, it is necessary to take into account the communication model that determines the probability that an opportunistic interaction occurs and, in this case, the distribution of the distance d between the nodes when the interaction takes place. The pdf $f_d(x)$ of the distance d, given by the derivative of (4.5)

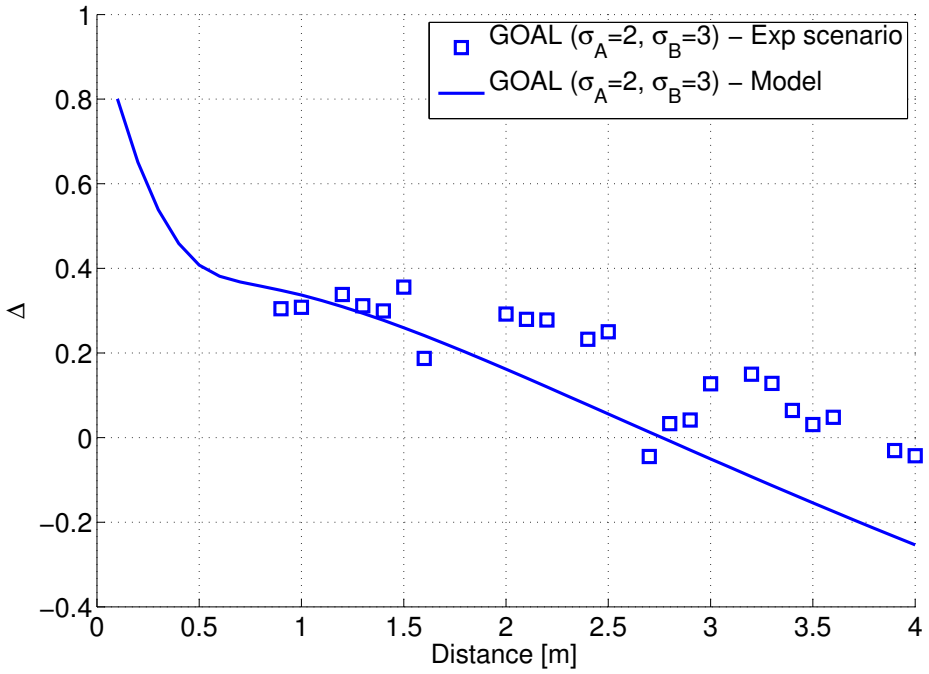


Figure 4.9: Theoretical vs experimental results for GOAL ($\sigma_A = 2$ m, $\sigma_B = 3$ m)

depends on different system parameters, such as the transmission range R, the duration T of the ON-OFF period that is usually implemented by wireless devices to save energy and the duty cycle δ, which is the fraction of the period T during which a node is available for opportunistic interactions.

Given $f_d(x)$, the final distribution of the positioning error is obtained by averaging the conditional distribution $F_{\tilde{\epsilon}_B}(x|d)$, given by (4.12), (4.14) or (4.18), over the distribution of d, which yields

$$F_{\tilde{\epsilon}_B}(x) = \int_0^R F_{\tilde{\epsilon}_B}(x|d)\, f_d(x) dx. \tag{4.21}$$

Fig. 4.12 reports the *opportunistic gains* Δ_B obtained for the heterogeneous scenario, when varying the duty cycle δ and the coverage range R. We observe that the best performance are obtained by setting $\delta = 0.5$ for every coverage range, though the performance improves for small value of R, as expected. This value is a good trade-off between the number of opportunistic updates and the distance at which they occur. It is interesting to see that, as described before, if node A has rather good self-localization capability and the ranging is reliable, GOAL performs slightly better than ML, with a gain of 30% over native positioning error, when the coverage range is $R = 2$ m. On the contrary, the ML approach is robust even if the range increases up to 4 m, where it reaches a 10% of gain, whereas GOAL's gain is limited to a coverage range of 3 m.

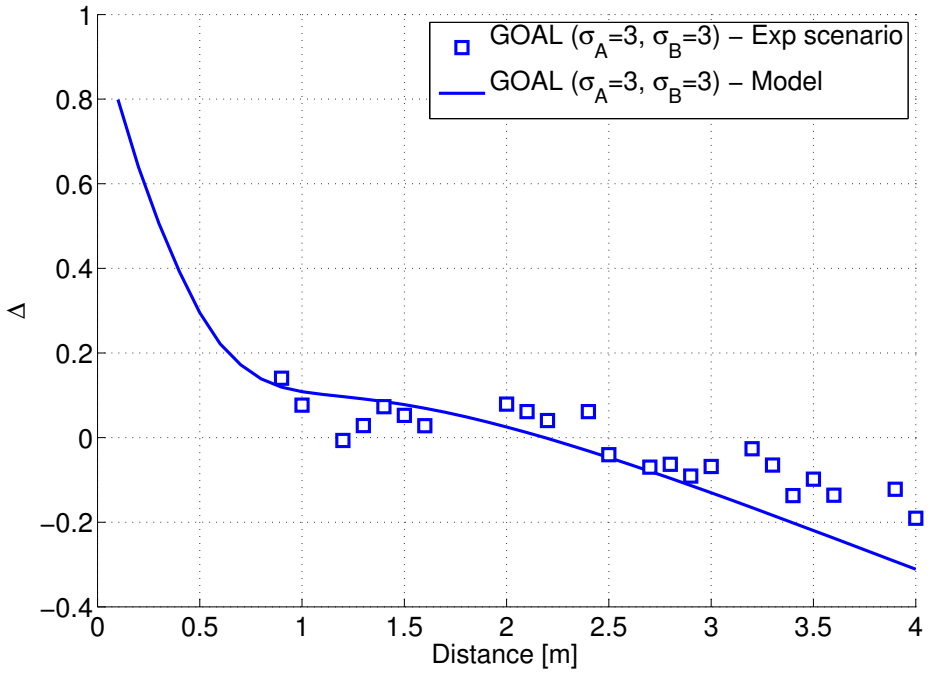


Figure 4.10: Theoretical vs experimental results of GOAL ($\sigma_A = 3$ m, $\sigma_B = 3$ m)

Fig. 4.13 shows the same results for the almost homogeneous scenario, in which nodes have similar (poor) native self-localization capabilities. The results for the ML scheme are substantially similar to those of the former case, though the curves are now more compacted and the relative error gain is reduced. This confirms the robustness of the statistical method. Also in this harsh scenario, GOAL performs as ML when the coverage range remains under 2 m, whereas performance rapidly decreases for larger coverage ranges. For $R \geq 4$ m, GOAL does not yield any gain.

4.3.5 Discussion

In this work we proposed and analyzed ML and GOAL, two localization techniques based on opportunistic data exchange between heterogeneous nodes. The ML algorithm has strong theoretical foundation, leveraging on the Maximum Likelihood paradigm, whereas GOAL leverages on a greedy, heuristic paradigm.

The algorithms have been investigated by using both theoretical analysis and real-world experiments. Results show that the opportunistic schemes can effectively improve the localization accuracy of the nodes, even though the gain closely depends on the settings of some system parameters. In particular, the duration of the scan period and its duty cycle, as well as the opportunistic transmission range need to be accurately set in order to attain

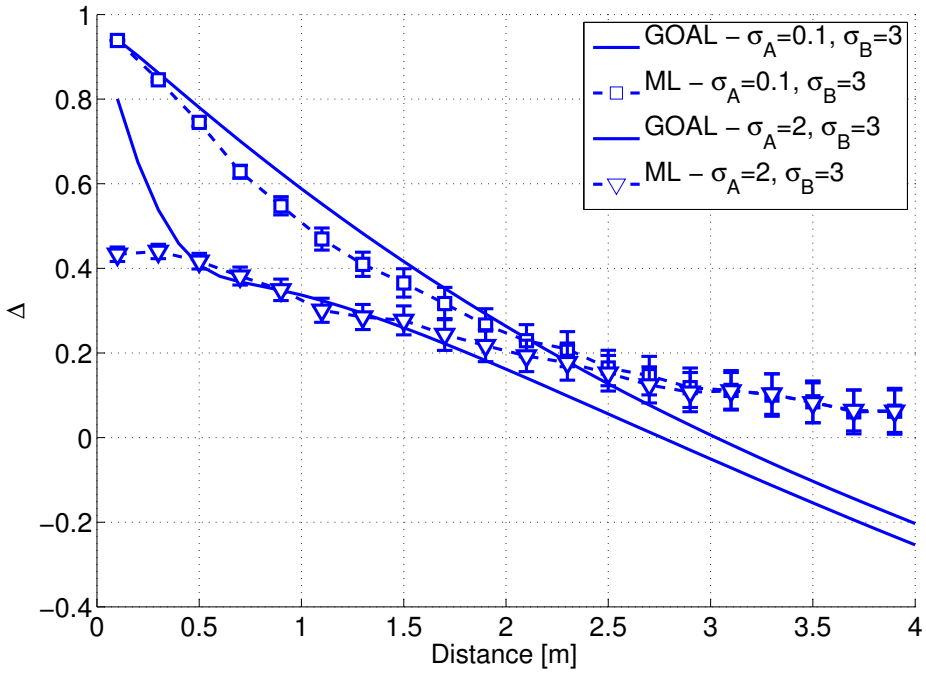


Figure 4.11: Relative localization error gain (4.20) after an opportunistic update at distance d.

significative gain.

GOAL, which is very simple, yields good performance only when one of the nodes is quite well localized and/or the interaction occurs at short distance. In this situation, GOAL yields $20 \div 30\%$ of positioning error gain over the native error when the opportunistic coverage range R is less than two meters. The ML approach is more reliable than GOAL and it yields performance gain even in quite harsh situations. The main drawback is the computational load that makes ML unfeasible for a real–time scenario using tiny devices as sensor nodes.

4.4 Linear Matrix Inequality

We address a scenario where a number of mobile nodes, equipped with a common communication device (Bluetooth, WiFi or ZigBee), can exchange data on an opportunistic basis, when they happen to be in coverage range. We suppose one node, called *User*, is not natively capable of localize its own position, whereas the other nodes, named *Peers*, can perform self-localization with a certain accuracy. The opportunistic localization problem consists in estimating User's position from the positioning information provided by the Peers through opportunistic data exchange.

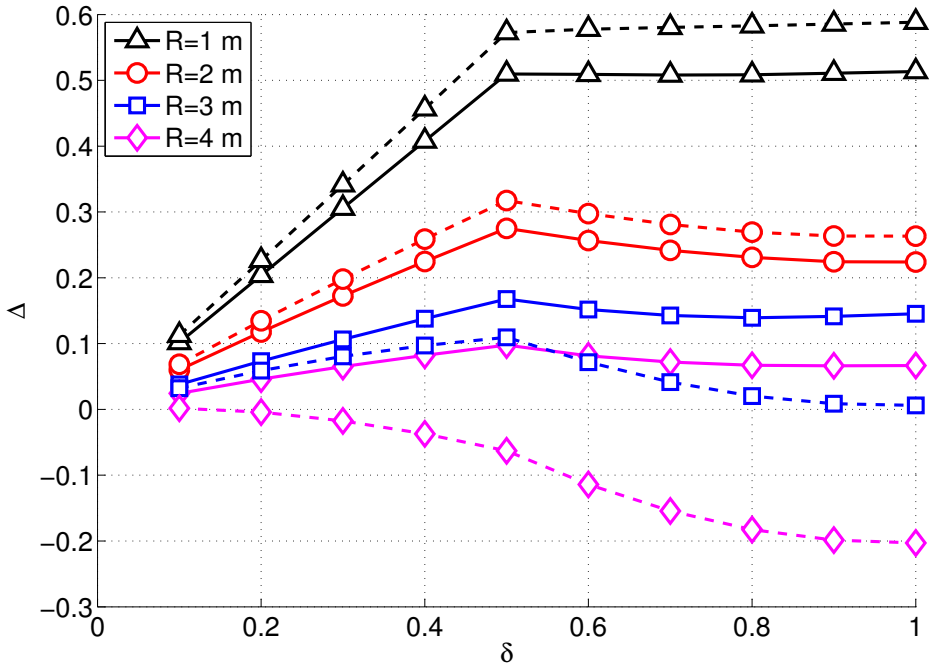


Figure 4.12: Relative localization error gain (4.20) after an opportunistic update for different values of range R in the scenario with $\sigma_A = 0.1$ m, $\sigma_B = 3$ m and $\sigma_\psi = 6$ dB. Solid lines refer to ML, dashed lines to GOAL.

4.4.1 Communication model

We assume every node in the network is equipped with a common wireless communication interface that is used for (opportunistic) data exchange. Radio propagation is modeled through the classic path loss and shadowing model as presented in (2.7)

We assume that opportunistic data exchange is enabled only if the received power is above a certain threshold $P_{th} \geq P_{min}$, where P_{min} is the minimum power level that guarantees successful packet decoding with high probability. The threshold P_{th} defines the so-called *nominal opportunistic range* R_{th} given by

$$R_{th} = d_0 10^{\frac{P_t + K - P_{th}}{10\eta}} . \tag{4.22}$$

In principle, opportunistic data exchange shall be enabled only between nodes within distance R_{th}. However, the random fluctuations of the received power due to the shadowing term ψ in (2.7) makes the actual range R_{opp} for opportunistic data exchange a log-normal distributed random variable, given by

$$R_{opp} = R_{th} 10^{\frac{\psi}{10\eta}} . \tag{4.23}$$

We observe that R_{opp} can take any (positive) value, though the probability of large values

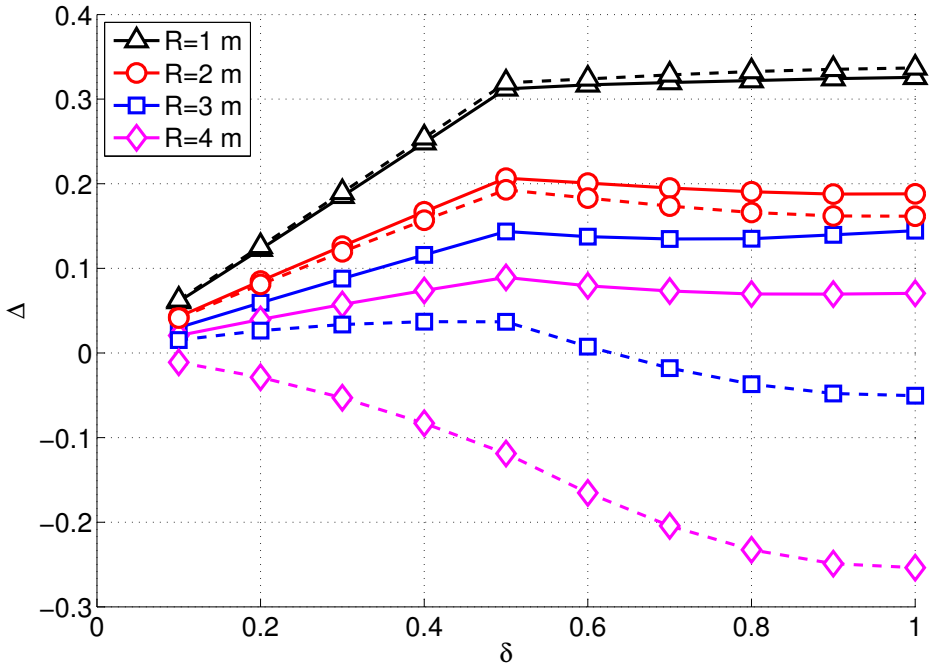


Figure 4.13: Relative localization error gain ((4.20)) after an opportunistic update for different values of range R in the scenario with $\sigma_A = 2$ m, $\sigma_B = 3$ m and $\sigma_\psi = 6$ dB. Solid lines refer to ML, dashed lines to GOAL.

goes rapidly to zero. Since the opportunistic localization algorithms described later in this paper require to fix an upper bound on the opportunistic range, we define R_{max} as the distance at which opportunistic data exchange occurs with probability 0.01, that is to say

$$\mathrm{P}[R_{opp} \leq R_{max}] = 0.99\,. \tag{4.24}$$

We hence assume nodes are uniformly distributed in a circle with radius R_{max} centered in the User and, according to (2.7), we admit opportunistic data exchange between Peer and User only when the received power at the User is above P_{th}.

4.4.2 Ranging model

We assume that, during an opportunistic data exchange, nodes can perform some sort of *ranging*, i.e., they can estimate their distance from some physical measurements, such as the received signal strength indication (RSSI) or the Time–of–Arrival of the radio signals. In our study, we consider two possible ranging scenarios, namely *range-free*, and *range-based*.

Range-free scenario

In this case, we consider a dummy ranging technique that provides a constant range estimate

$$\hat{d} = R_{th}\,, \tag{4.25}$$

irrespective of the actual distance between transmitter and receiver (provided that the received power is above the opportunistic threshold P_{th}). The maximum ranging error is hence equal to

$$\epsilon_{rang}^{max} = R_{max} - R_{th}\,, \tag{4.26}$$

with high probability. In practice, localization algorithms based on binary ranging are in all respects range-free algorithms.

Range-based scenario

In this case we assume an RSSI–based ranging technique, which uses the value of signal power Γ_r measured by the receiver to provide a Maximum-likelihood estimate $\hat{d}(\Gamma_r)$ of the real distance d between transmitter and receiver. The estimate is obtained by reversing the path-loss equation (2.7) once neglected the shadowing term as in (2.9).As for the opportunistic range, we also define an upper bound d_{max} on the RSSI-based ranging estimate, such that

$$\mathrm{P}[\hat{d} \leq d_{max}] < 0.99\,. \tag{4.27}$$

Hence, the ranging error for a range estimate $\hat{d}$ with high probability is limited to

$$\epsilon_{rang}^{max} = d_{max} - \hat{d}\,. \tag{4.28}$$

We observe that, setting the shadowing variance σ_ψ to zero, this ranging technique provides perfect distance estimation. Although this condition is not feasible in practice, it is interesting to gain insight on the impact of ranging on opportunistic localization algorithms performance.

4.4.3 Self-positioning model used by peers

We assume that peer nodes have "native" self-positioning capabilities, provided by some (non opportunistic) scheme. Accordingly, we denote by $\mathbf{s}_i$ and $\hat{\mathbf{s}}_i$ the real and the self-estimated position of the ith peer. For simplicity, we assume that the estimation error $e_i = \|\mathbf{s}_i - \hat{\mathbf{s}}_i\|$ can be modeled as the module of a 2–dimensional Gaussian Random Variable $[x,\ y]$, with zero mean and variance σ_{loc}^2. Accordingly, the self-localization *distance* error $\epsilon = \sqrt{x^2 + y^2}$ is a Rayleigh-distributed random variable with parameter σ_{loc}. Hence, the parameter σ_{loc} defines the accuracy of the native self-localization mechanism supported by the node. For

simplicity, in each simulation we assume that all Peers have the same σ_{loc}. Once again, it is practical to set an upper bound ϵ_{loc}^{max} on the self–localization error of Peers, such that

$$\mathrm{P}[\epsilon \leq \epsilon_{loc}^{max}] = 0.95\,. \tag{4.29}$$

During an opportunistic data exchange, Peers send packets containing their estimated positions $\hat{\mathbf{s}}_i$ and their class of accuracy σ_{loc}^2. This information, together with the ranging estimate, is then used by User to estimate its own position through the opportunistic localization mechanisms described below.

4.4.4 Opportunistic localization algorithms

As mentioned, User resorts to opportunistic localization to infer its geographical position. To this end, User opportunistically collects information from passing by Peers. Information consists of i) $\hat{\mathbf{s}}_i$, which is the current position estimate of Peer i, provided by self-localization system; ii) ϵ_{loc}^{max}, which is the maximum self-localization error, as given by (4.29); iii) a ranging estimate $\hat{d}$ and maximum ranging error ϵ_{rang}^{max}, which depends on the ranging technique adopted by User, as specified in Sec 4.4.2. User is assumed to remain still while collecting information for opportunistic localization. Then, it performs one of the opportunistic localization algorithms described below to get an estimate $\hat{\mathbf{s}}_u$ of its real position $\mathbf{s}_u$.

Centroid algorithm

The Centroid algorithm estimates User's position simply as

$$\hat{\mathbf{s}}_u = \sum_{i=1}^{N} \hat{\mathbf{s}}_i\,, \tag{4.30}$$

where N is the total number of Peers encountered during the opportunistic localization phase. An example of how Centroid algorithm works is illustrated in Fig. 4.14.

Centroid does not rely upon any ranging information, but only on Peers self–localization estimations. If Peers are uniformly distributed around User, then Centroid will asymptotically provide the correct User position. This scheme is really simple, not requiring any knowledge about the channel and self-positioning error models. On the other hand, its performance strongly depends on the spatial distribution of Peers. Although this localization algorithm is very basic, it provides a useful benchmark to compare the performance of more advanced schemes.

LMI localization algorithm

The main approach for opportunistic localization considered in this paper is based on the solution of Linear Matrix Inequality (LMI) problems. Basically, for each Peer i involved in

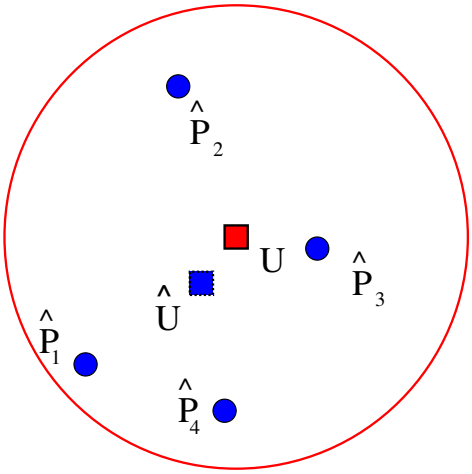


Figure 4.14: Pure barycenter estimation.

opportunistic data exchange, User writes the following inequality

$$\|\mathbf{s}_u - \hat{\mathbf{s}}_i\| \leq R_i \tag{4.31}$$

where $||\cdot||$ denotes the euclidian distance, and R is the maximum admissible distance between User and Peer i given by

$$R_i = \hat{d} + \epsilon_{rang}^{max} + \epsilon_{loc}^{max} \,. \tag{4.32}$$

The set of inequalities (4.31) collected during the opportunistic phase confines User's position in the area covered by the intersection of the circles centered in $\hat{\mathbf{s}}_i$ and having radius R_i, for $i = 1, 2, \ldots, N$. The User's position is then estimated by choosing a point $\hat{\mathbf{s}}_u$ within this area, possible at maximum distance from the area border. This problem can be formulated as a Linear Matrix Inequality (LMI) problem [44], which can be solved with standard techniques.[1] Fig. 4.15 exemplifies the idea in the case of four contacts with Peers P_1 to P_4.

We observe that, due to the bounding (4.24), (4.27), and (4.29), it is actually possible that User receives messages from Peers farther than R_i, though with very low probability. In this case, however, the circles intersection is empty and the LMI problem becomes unfeasible. When this event happens, we solve again the LMI problem increasing the value of R_i. In the range–based algorithm, we consider a new value of estimated distance $\hat{d}'$, obtained multiplying $\hat{d}$ by a magnifying factor $m > 1$ ($m = 1.2$ in our experiments). Instead, in the range–free algorithm, because R_i is the same for each contact, we increase it with a constant, using in our experiments $R_i + 1$. The procedure is repeated until the problem becomes feasible.

[1] For instance, MATLAB solver `mincx` gives as an output the coordinates of the rectangle circumscribing the circles intersection. The center of the rectangle may then be chosen as LMI solution.

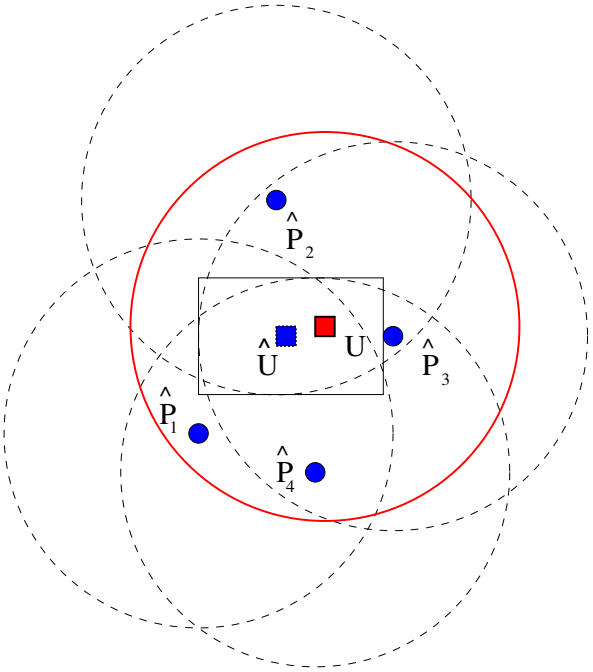


Figure 4.15: Pure LMI estimation.

Γ_t	η	K	d_0
0 dBm	1.74	-36 dB	0.1 m

Table 4.2: Path–loss channel parameters

Hybrid LMI-Centroid algorithm

Although there are efficient numerical methods to solve an LMI, the resolution time rapidly grows with the number of inequalities in the system. To cope with this problem, we propose a hybrid algorithm that combines the LMI and Centroid techniques. Basically, to limit the computational cost of LMI we partition the N inequalities in $K = N/L$ groups of L inequalities, thus obtaining K temporary estimates $\hat{\mathbf{s}}_{u,r}(k)$, with $k = 1, 2, \ldots, K$. Then, we apply the barycentric technique to obtain the final estimate $\hat{\mathbf{s}}_u$ as

$$\hat{\mathbf{s}}_u \quad = \quad \sum_{k=1}^{K} \hat{\mathbf{s}}_{u,r}(k)\,. \tag{4.33}$$

The hybrid scheme makes it possible to trade off localization accuracy for computational cost. Fig. 4.16 exemplifies the result of this technique in the same scenario considered for the previous algorithms.

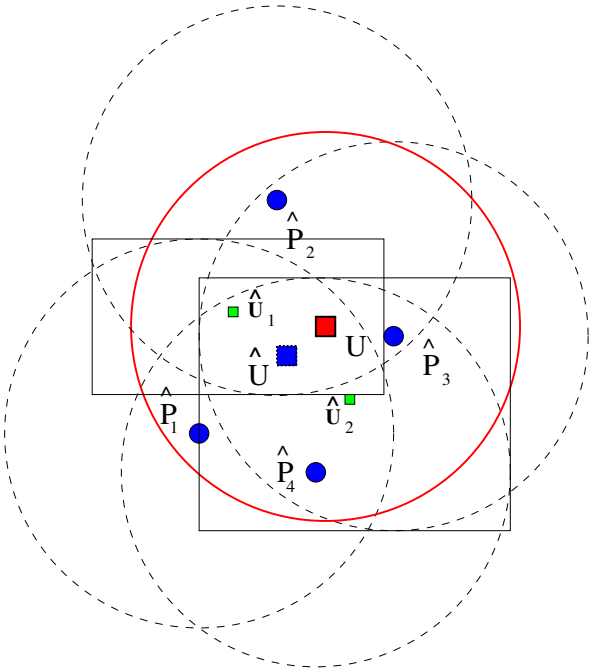


Figure 4.16: HLMI estimation with L=2.

4.4.5 Performance analysis

In this section, we analyze the results provided by the three opportunistic localization algorithms, namely Centroid, LMI and Hybrid LMI, in different scenarios. For all the experiments, we consider the path-loss channel parameters reported in Table 4.2 and $R_{th} = 10$ meters. For each scenario, we ran 60 experiments to obtain enough statistical accuracy. We calculated the confidence intervals for each scenario, having less than 10 cm in a stable scenario and less than 25 cm when adding sources of error.

We first compare the LMI and Centroid algorithms, considering both the range–free and range–based case for LMI. To begin with, we study the performance of the algorithms in an ideal scenario, where Peers have perfect self–localization ($\sigma_{loc} = 0$) and the radio channel is deterministic ($\sigma_{\psi} = 0$). Hence, opportunistic data exchange only occurs within a range R_{th} from User, and ranging provided by (2.9) is error-free. As expected, range–based LMI largely outperforms its range–free counterpart that, in turn, performs better than Centroid. Range–based LMI achieves very accurate localization already after 10 opportunistic contacts, whereas range–free LMI and Centroid algorithms reach low localization errors after a larger number of contacts. These results confirm that, against a higher complexity, the LMI approach is actually capable of yielding a significant performance enhancement over the simple Centroid algorithm, at least in ideal conditions. The question now is whether or not such a performance gain will be maintained under more realistic assumptions.

In Fig. 4.17, we report the results obtained for an ideal radio channel model ($\sigma_{\psi} = 0$),

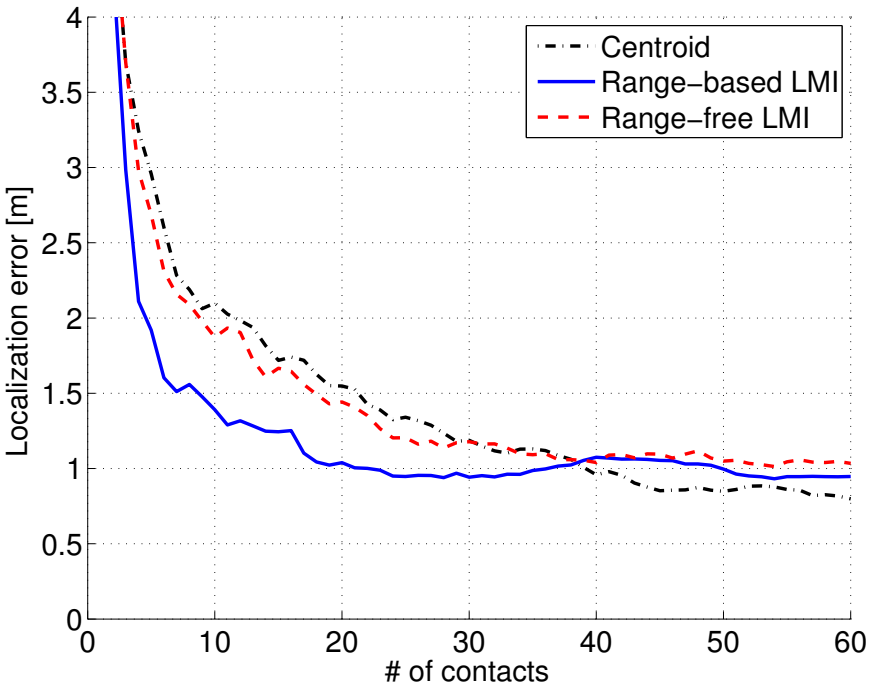


Figure 4.17: Localization performance of range–free LMI, range–based LMI and centroid algorithms, with $\sigma_\psi = 0$ dB and $\sigma_{loc} = 2$ m.

but assuming that Peers have native self-localization error with variance $\sigma_{loc} = 2$ m. As expected, the performance of LMI algorithms worsens with respect to the ideal case, whereas the Centroid approach seems to be unaffected by Peers' self-localization errors. This is due to the fact that self-localization errors of Peers are assumed to have zero mean, so that if Peers are uniformly distributed in the area, after a sufficient number of contacts Centroid will still return an estimate close to User's real position. With perfect ranging and Peers' self-localization error with $\sigma_{loc} = 2$ m, the range-based LMI still exhibits better performance than the other algorithms, whereas range-free LMI and Centroid performance are comparable. However, the LMI algorithms suffer severely if Peers' self-localization error increases ($\sigma_{loc} = 4$ m), in which case performance drops below Centroid's one.

In Fig. 4.18, we analyze the behavior of the algorithms in case of perfect Peers' self-localization of Peers ($\sigma_{loc} = 0$) and signal propagation affect by shadowing with $\sigma_\psi = 2$ dB. We note that results are, in general, worse than for the previous case. The randomness of the signal propagation, in fact, brings along ranging errors and unpredictable opportunistic data exchange range, since communication can occur even at distances larger than R_{th}, whereas it may happen that nodes within R_{th} do not communicate. With $\sigma_\psi = 2$, range–based LMI performs better than Centroid, whereas for larger value of σ_{psi}, e.g., with $\sigma_\psi = 4$, LMI has quite poor accuracy (the y-axis scale is larger than the other scenarios), being above

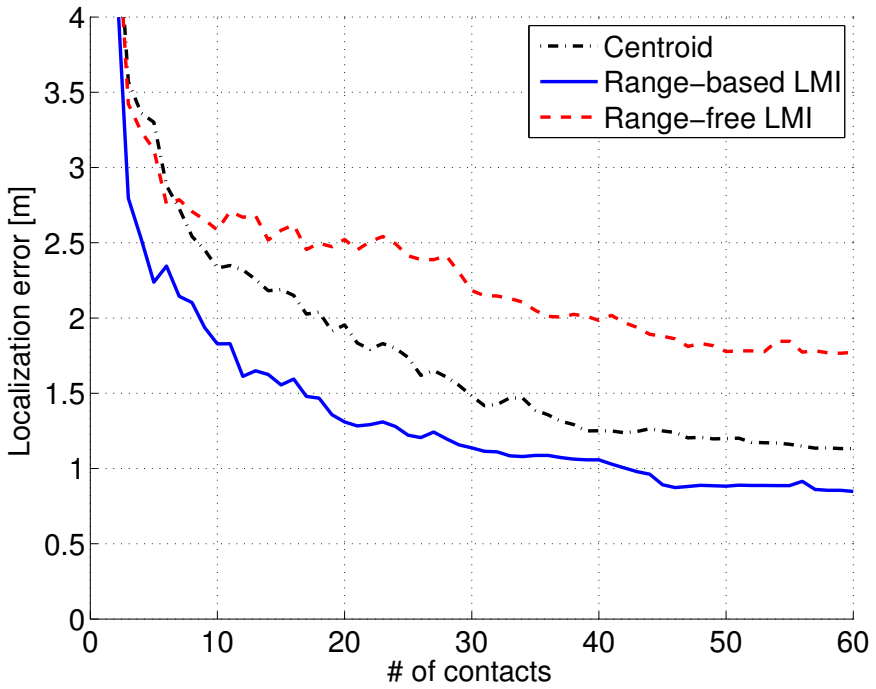


Figure 4.18: Localization performance of range–free LMI, range–based LMI and centroid algorithms, with $\sigma_\psi = 2$ dB and $\sigma_{loc} = 0$ m.

Centroid only for the first 5 contacts. In this scenario, for a larger number of contacts, the probability that the LMI problem becomes unfeasible grows, so that it is necessary to relax the constraints as explained in Sec. 4.4.4, thus worsening performance. It is interesting to note that, somehow unexpectedly, range–free LMI performance are worse than range-based LMI's one, despite the unreliable ranging. The reason is that, according to (4.32), the range-free LMI draws circles of rather large radius R_i, irrespective of the distance between Peer and User, so that the intersection area that contains User's position is rather wide. Instead, since the ranging error is statistically smaller when the distance between nodes is short, then the range-based LMI draws circles of smaller radius around the estimated positions of Peers that are likely to be close by User. These circles shrinks the intersection area around User position, thus increasing the estimate accuracy.

In summary, the LMI approach, in particular the range-based, is potentially capable of yielding good localization, though it suffers from high computational complexity and performance degradation for large number of contacts and unreliable channels. In the following we hence investigate the performance of the hybrid LMI approach, which may alleviate some of these problems.

In Fig. 4.19, the three red solid lines show the performance of Hybrid LMI with range–free technique, for different values of L, with deterministic radio channel and unreliable Peers'

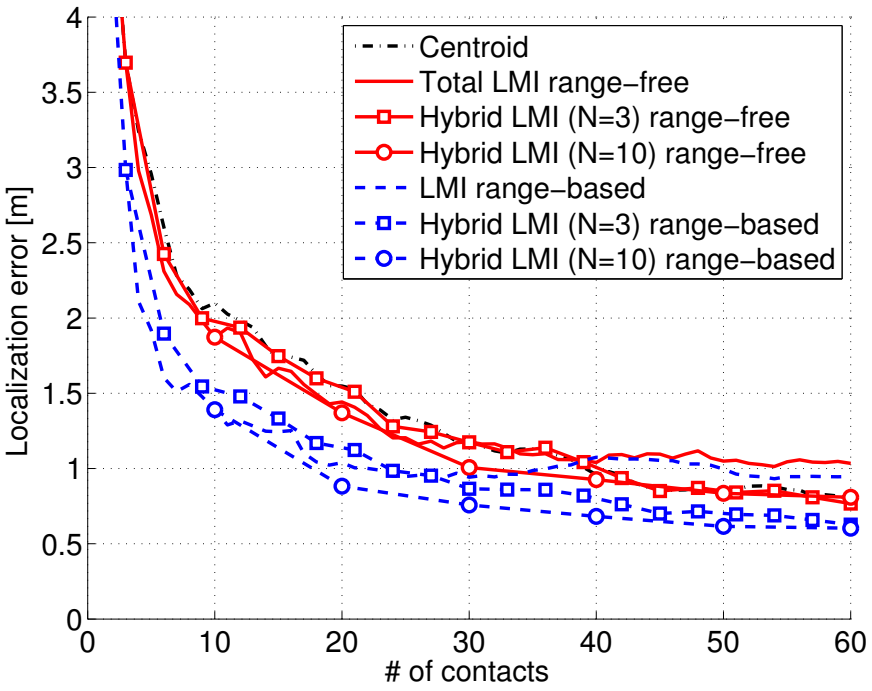


Figure 4.19: Localization performance comparison between range–free and range–based Hybrid LMI with different values of L, with $\sigma_\psi = 0$ dB, $\sigma_{loc} = 2$ m.

self-localization. First, it is possible to notice that, with $L = 3$, the Hybrid algorithm achieves basically the same performance of Centroid. With $L = 10$ the Hybrid LMI performs slightly better than Centroid. The blue dashed lines in Fig. 4.19 show the same results for range–based Hybrid LMI scheme. We see that, also in this case, the Hybrid approach yields better performance than pure-based approach when the number of contacts increases, in particular when self-localization error of Peers increases.

When channel conditions become harsh, the problem of unfeasibility of the LMI solution for large numbers of inequalities has a strong impact. Therefore it is more convenient to use a small L, both in range–free and range–based LMI to achieve better results, as shown in Fig. 4.20. Where the classic LMI range–based has worse performance than the Centroid, the HLMI achieves better performance than Centroid, and in addition it also decreases the computational complexity. The same behavior can be seen for the range–free HLMI: the less the L, the more the accuracy, approaching very close to the Centroid performance.

4.4.6 Discussion

In this work, we addressed the problem of the localization of stray node, called User, by means of opportunistic data exchange with passing by nodes, called Peers, which are instead capable of self-localization. We studied two main localization schemes, namely Centroid, in

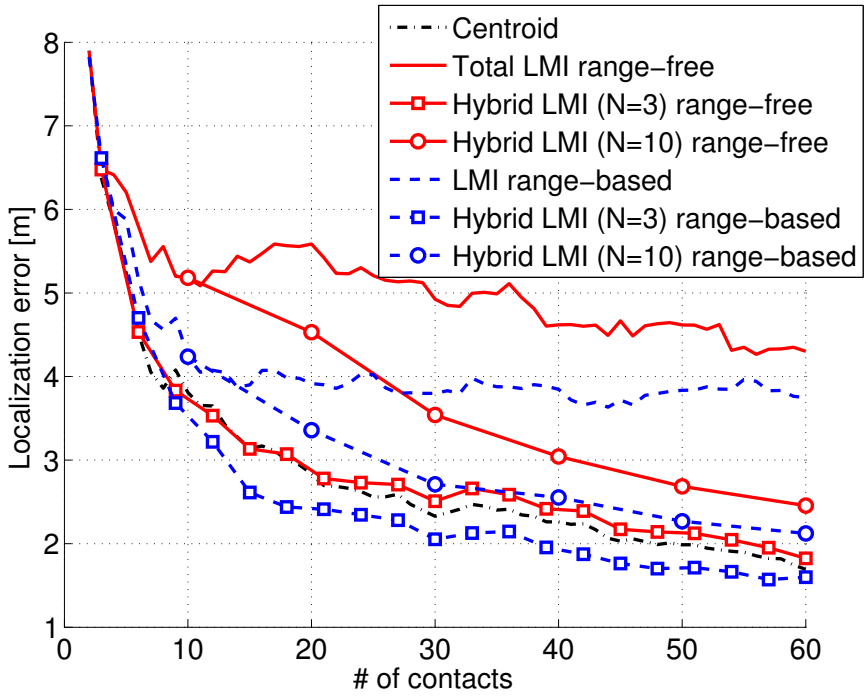


Figure 4.20: Localization performance comparison between range–free and range–based Hybrid LMI with different values of L, with $\sigma_\psi = 4$ dB, $\sigma_{loc} = 2$ m.

which User estimates its position as average of Peers' position estimations, and LMI that takes into account a connectivity model and a ranging model to provide a bound on the distance between Peer and User.

We simulated different scenarios, varying the randomness of the signal propagation and the self-localization accuracy of Peers. Results revealed that, in favorable scenarios, the LMI localization achieves very good results, especially for the range–based scheme. On the contrary, when the signal propagation is random or Peers self–localization is not reliable, LMI performance degrades faster than Centroid's one. Furthermore, we observed that the LMI localization performance worsens, compared to the Centroid technique, when the number of contacts exceeds a given threshold with unreliable channels, because the constraints considered in the LMI problem becomes unfeasible and need to be relaxed. Hence, we analyzed the Hybrid LMI algorithm that splits the LMI problem in subproblems of L inequalities that are resolved independently and, then, computes the barycenter of the different solutions. In this case, localization accuracy strongly depends on the choice of L, though a suitable tuning of this parameter is effective in counteracting the performance loss due to constraints relaxation. Furthermore, Hybrid LMI algorithm dramatically reduces the computational effort.

Chapter 5

Traffic management in radio wireless sensor network

A common scenario in WSNs, which occurs in many applications, considers a "convergecasting" communication paradigm. In this case, the whole network strives to harvest data and have them reported to the sink, which is assumed to be in charge of any collection and processing. While this simplifies the task of the sensors, it has a clear disadvantage, since all traffic gathers around the sink. As a consequence, communications originated by different neighbors are likely to collide due to channel congestion and, in particularly high traffic conditions, severe data losses may be experienced. In addition, this congestion may spread to the rest of the network, resulting in a generalized deadlock.

The chosen communication protocol and the general network design play a major role in the congestion avoidance properties of the WSN. For instance, choosing an 802.11 DCF-like [58] Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) protocol results in simple design but poor efficiency. In fact, Request–To–Send (RTS) and Clear–To–Send (CTS) messages used for silencing the neighbors of two parties engaged in a communication tend to decrease parallelism, as two links cannot coexist in the same area. Since the access delay of a contention-based protocol is generally larger than the delay of a deterministic protocol, this may reduce the amount of traffic that reaches the sink.

5.1 State of the art

The problem of congestion near the sink in a WSN has recently received increased attention, as convergecasting solutions and related protocol research are gaining momentum. Some techniques and protocols that try to mitigate and solve this problem have been proposed. The topic is interesting not only from the point of view of traffic and reliability in the transmission, but also from an energy point of view. In fact, congestion means also a lot of

wasted energy due to retransmissions and multiple channel access attempts.

In CODA (COngestion Detection and Avoidance) [59], the authors propose two solutions: an open-loop hop-by-hop backpressure and a closed-loop multi-source regulation. In the first case, a node broadcasts a backpressure message to its neighbors as soon as it detects congestion. The nodes that receive this message adapt their sending rates, or drop packets according to a certain local policy. The main drawback here is that, in the presence of strong deadlock, it could be very difficult to broadcast the backpressure message. The second proposed solution consists of a sort of end-to-end congestion control, enforcing cooperation between the sources and the sink. Specifically, when a source detects an amount of traffic exceeding a fraction of the theoretical channel throughput, it sends packets with a notification bit that forces the sink to reply with acknowledgment messages (ACKs). Hence, if no ACKs reach the source, congestion is detected. However, an end-to-end solution may be more costly, especially if congestion prevents notifications from reaching the sink.

A similar approach to congestion control, Event-to-Sink Reliable Transport (ESRT), can be found in [60]. This protocol lets nodes independently check their queue level. When it exceeds a certain threshold, the source marks outgoing packets with a notification bit. When the sink receives one such packet, it sends a message to request all sources to decrease their sending rate, so that congestion levels fade and normal network operations can be restored. The message sent by the sink may be transmitted at high power, in order to reach the greatest number of nodes, but in turn this creates interference that may hamper currently on-going handshakes.

In [61], the authors propose a technique to modulate the traffic generation rate. Basically, they construct a tree-based topology where every node knows the number of its children nodes and the traffic rate of its parent. It then determines its own sending rate and communicates it to its children, thereby implementing a sort of backpressure. The purpose is to equally distribute traffic and channel access in order to avoid congestion and enforce fairness, especially for the packets that need to cross multiple hops for reaching the sink. A drawback of this work is that the authors assume that all sensors are always on, whereas in a WSN it would be desirable to employ duty cycling for the sake of energy saving.

In [62], the authors aim at assessing how local parameters can globally affect network congestion. Their study shows how it is possible to achieve good results by varying simple parameters, such as the number of sources, the buffer size and the retransmission timeout.

The approach we follow in our work is in part similar to the existing literature, in that we try to mitigate near-sink congestion. On the other hand, it is also different, as we assume that the neighbors of the sink can afford to behave differently than all other sensors in order to gain some throughput advantage. Our point is whether it is more convenient to enhance MAC strategies near the sink and allow for two different protocols (one executed near the sink, the other in the rest of the network) to co-exist and be interfaced, or if it is better to

use and optimize a single protocol.

5.2 EPC^2

In this section, we explore a different approach to the congestion problem. Namely, we assume that the nodes placed near the sink can detect that they are in fact its one-hop neighbors, and configure themselves to communicate with the sink using a special protocol, that will be specified later. This way, two problems arise: *i*) the protocol must be properly designed, in order to leave as little traffic as possible unserved, and *ii*) it must be interfaced with the other protocol used by the rest of the network until one-hop neighbors of the sink are reached. Our main point here is whether it is more convenient to resort to an efficient (but perhaps needing more design effort) two-protocol solution, or if it is better to hinge on a single protocol.

After having introduced our approach, we show the impact of some protocol parameters on performance, and compare it to plain CSMA/CA. Our comparison includes an event-based traffic model, meant to represent a more realistic networking scenario. Ultimately, we show that there is in fact some benefit to be gained by leveraging on a solution that allows more traffic to get through near the sink, even if this means a greater design effort for interfacing two protocols which must coexist in the same network.

5.2.1 Protocol Description

Carrier Sense Multiple Access with Collision Avoidance

CSMA is a simple, completely distributed approach for accessing a shared channel. It does not require any synchronization or coordination among nodes, and thus may be suitable to resource constrained WSNs. With CSMA, a node can send packets only if it senses a free channel for a certain amount of time, otherwise it backs off and reschedules a later attempt. Nonetheless, such solution suffers from the hidden terminal problem [63]. The CA feature adds some robustness in this sense. Prior to data transmission, the sender issues a Request–to–Send (RTS) message, which alerts the receiver and silences all other neighbors of the transmitter. A Clear–to–Send (CTS) packet is then sent by the addressed receiver to confirm its availability and to silence the receiver's neighbors as well, thus preventing hidden terminals to interfere with the communication. On one hand, this protocols yields reasonable performance in terms of throughput and packet transmission success, if the overall network traffic is not too high. On the other hand, the preemptive signaling is an inherent communication overhead. Moreover, in dense networks and under high traffic conditions, blocking nodes for avoiding collisions may prove to dramatically reduce the communication parallelism and to increase the number of nodes in backoff.

Efficient Packet Converge-Casting (EPC^2)

Our novel approach consists in binding and interfacing simple CSMA/CA operation throughout the network with a novel protocol run only by the sink and its neighbors. This protocol is dubbed Efficient Packet ConvergeCasting (EPC^2), because it aims at improving the efficiency of packet delivery near the sink, so that all requests by all neighbors are accomodated, thereby increasing the local throughput and transmission success ratio, and translating these advantages into an overall greater forwarding effectiveness. The protocol is partially inspired to slotted time–division multiple access.[1]

With EPC^2, the one-hop neighbors of the sink operate in cycles made of 4 phases.

Receive phase: the neighbors only receive packets from the rest of the network. Their behavior follows a CSMA/CA scheme. The duration of this phase is set and communicated by the sink in an adaptive way, according to specific metrics that will be described later.

Guard phase: this phase is necessary in order to avoid overlap between the receive phase and the subsequent notify phase. In the worst case, an RTS may be received by a neighbor of the sink just a little before the receiving phase ends, therefore the guard phase has to accomodate enough time for a whole data transmission (one RTS, two CTSs,[2] one data and one ACK transmission) to be completed.

Notify phase: during this phase, every node sends a small packet to the sink to communicate the length of its own backlog. According to the received information, the sink sets up a round-robin schedule of the transmissions of its one-hop neighbors, so that each node has enough time to empty its own queue. In order to keep the synchronization among the nodes, each packet transmission fits within a *slot* of length T_{data}. The slots assigned to each node are consecutive. The scheduling decisions taken by the sink are then broadcasted to the neighbors, along with the duration of the next receive phase. If this packet is lost, the node is not allowed to transmit.

Transmit phase: Each neighbor sends its packets to the sink, according to the received schedule. The nodes stay asleep during others' transmissions and wake up only for their turn. If a node does not receive the schedule, it stays asleep for the whole duration of the transmit phase.

In Figure 5.1, the operation of EPC^2 is exemplified. N and A denote Notify and ACK packets, respectively. Notice that node 3 loses the schedule packet, and thus refrains from

[1] Note that such a technique requires synchronization among the nodes. However, as long as only the neighbors of the sink run EPC^2, this is easy to achieve, *e.g.*, by having the sink broadcast periodic beacons.

[2] One CTS time is needed for routing purposes, see below.

any transmission.

How to determine the length of the receive phase at the beginning of the next cycle is a design choice where some tradeoffs can be exploited. For example, on the one hand the choice affects the delivery delay, because if the receive phase is long, the communication latencies toward the sink grow accordingly. On the other hand, a longer receive phase allows the neighbors of the sink to receive more packets from the network, and consequently to deliver all of them (or nearly so) within a single handshake, thereby requiring fewer control messages and increasing efficiency. Nonetheless, more packets in the queue result in a longer transmit phase. Therefore, two-hop neighbors of the sink cannot manage to contact one-hop neighbors for a long time. Since they run a simple CSMA/CA access mechanism, this may force them to back off more frequently, or even to drop some packets. Thus, we leave the number of slots per node as a protocol parameter, M. Based on that, the sink can determine how many packets it should receive during the transmit phase, *i.e.*, MN_s, where N_s is the number of its neighbors. If the number of received packets is smaller than MN_s, the receive phase duration is increased by 0.1 s,[3] to let the neighbors gather more data packets before the transmit phase. Otherwise, it is reduced by the same amount of time. Other, more refined adaptive methods can apply as well, but we focus on this simple policy which has proven to give good results in our simulations.

Using a different protocol for the only neighbors of the sink introduces further complexity. First, a synchronization problem, since the nodes have to share a common TDMA schedule to access the channel correctly. Second, a protocol interfacing problem, whereby a special phase is needed for the neighbors of the sink to receive traffic from the rest of the network. Finally, the TDMA schedule and the transmission turns set up among the neighbors are expected to yield some further delay. This last problem will be assessed in detail in Section 5.2.2.

Since the network is not completely connected, reaching the sink needs multi-hop communications. In order to focus as much as possible on the effects of the MAC protocols being compared, we keep the routing mechanism very simple. Namely, we suppose that each node knows its own hop count (HC). In a static network, this can be accomplished by letting the sink transmit a beacon during an initialization phase: each node that received the beacon sets its own HC to 1, stores this value into the message and rebroadcasts it. The process goes on this way, with each node assuming as its own HC the minimum HC value read from a received packet, plus 1. Upon initiating a handshake, the transmitter inserts its HC in the RTS. Only nodes with an HC smaller than or equal to the transmitter's are allowed to answer with a CTS. The CTS is issued after a random backoff time, uniformly chosen between 0 and one CTS time. This backoff is necessary to make collisions among CTSs less likely, and

[3]The value of this time difference can be set to adapt the protocol response speed and to trade it off with stability in the number of received packets. After a thorough study, the value 0.1 s proved to work well in our simulations.

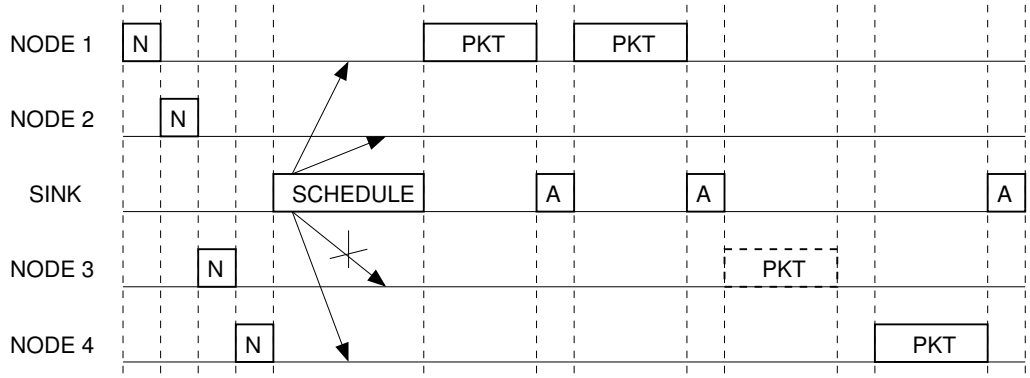


Figure 5.1: Example of EPC^2 signaling and transmissions.

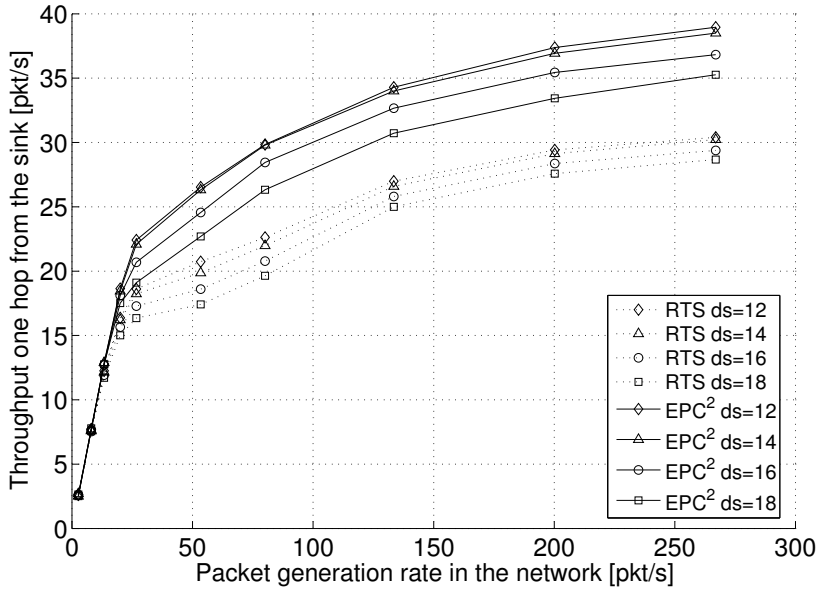


Figure 5.2: Throughput (last hop).

to ultimately choose one relay that will receive the data packet. This also explains why two CTS times are included in the duration of the guard phase.

5.2.2 Simulation Results

In this Section, we provide a thorough simulation study of our protocol, aimed at assessing its performance. We assume that the network is deployed within a circular area with radius $r = 25\,\mathrm{m}$, where 80 nodes are randomly positioned.

The sink is placed at the center of the area.

Each node is equipped with a radio that allows for a data rate of 33.6 Kbps using a standard BPSK digital modulation [64]. The maximum transmit power allowed by the radio

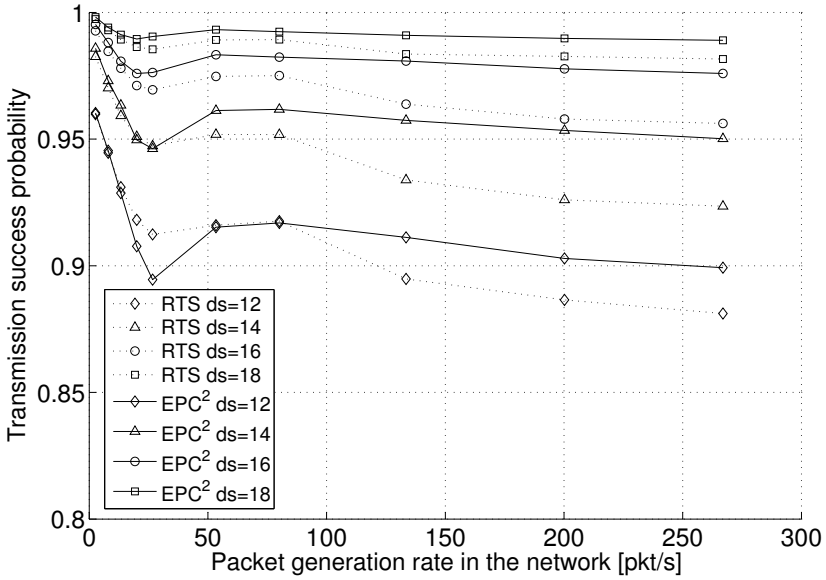


Figure 5.3: Transmission success ratio.

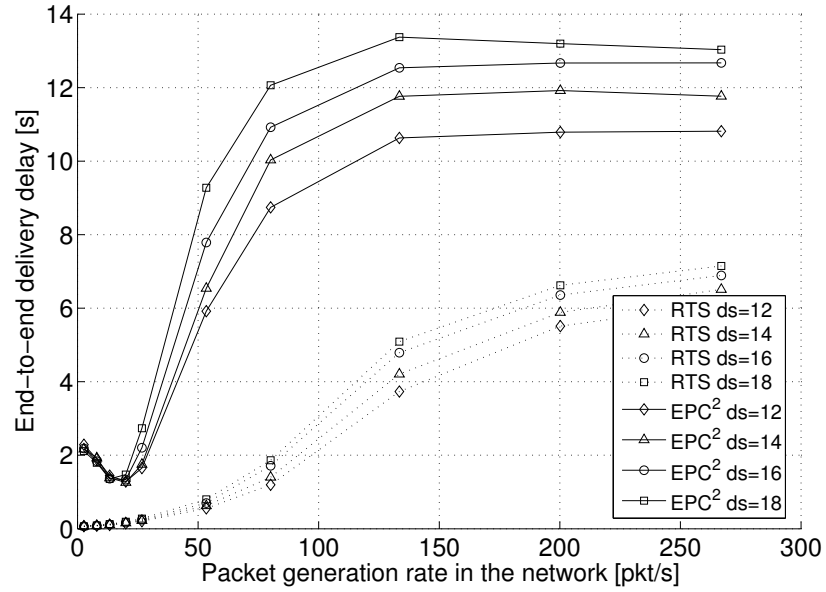


Figure 5.4: End-to-end delivery delay.

is 10 mW. We suppose that the power of a transmitted signal fades with increasing distance according to the Hata model [64], whereby the power loss is calculated as Ad^{β}, where d is the distance between the terminals, $\beta \geq 2$ is the path loss parameter, and $A = 10^3$ for transmissions employing a carrier frequency around 900 MHz. We choose $\beta = 3.5$ in order to model a harsh environment (*e.g.*, indoor or under foliage). The data packet length is fixed to be 200 bits, whereas any signaling message is 20 bits long. We assume that channel errors occur independently on each bit, so that the probability that a packet correctly reaches its recipient depends directly on the Signal-to-Interference-and-Noise-Ratio (SINR). Specifically, this probability is found to be almost 1 up to a distance $\bar{d} = 9$ m, according to the previous parameters. Therefore, $\bar{d}$ can be taken as an approximation of the node coverage range (in the sense that within d, transmissions are received correctly with high probability). It is worth highlighting that accounting for interference to take part in the calculation of the SINR allows to explicitly model the capture effect, whereby a packet could be correctly received even in the presence of some superimposed transmissions, provided that the SINR is sufficiently high.

We take into consideration two different traffic situations. In the first case, packets are generated according to a Poisson process with rate λ packets per second. Each packet is then randomly assigned a source. This scenario allows to assess the general performance of the protocols, and may represent a typical WSN configuration, where nodes are asked to sense some data and periodically report back. Therefore, we perform protocol parameter optimization in this scenario. In the second case, we simulate the protocol behavior with event-driven traffic, *i.e.*, by generating an event in a random position within the network area. Namely, the nodes generate traffic according to a Poisson process with low rate (0.1 pkt/s). At some time, an event is generated and sensed by all nodes within a range randomly chosen in the interval $[2, 16]$ m. Upon perceiving the event, each of them begins to generate data packets at a much higher rate (a parameter in our simulations). When the event stops, the network restores its low rate packet generation. In both scenarios, we assume that the neighbors of the sink do not generate data.

Since energy saving is a major concern for WSNs, we let nodes alternate between awake and sleep states according to a prescribed duty cycle. Extensive simulations showed that a reasonable value that yields good performance yet allowing for substantial savings is 20%. Moreover, every packet can be retransmitted up to a maximum of 20 attempts, after which it is dropped.

We simulated both protocols using MATLAB. The simulation time and the number of network realizations have been set so as to reach a sufficient statistical confidence. The rest of this Section is devoted to presenting simulation results.

Tuning EPC^2: the sensing range

First of all, we show here how the overall network performance is affected by the sensing range d_s of CSMA. The sensing range is determined by a threshold on the minimum received power that makes a node assume the channel busy. Therefore, it affects the average SINR experienced by a transmission: a greater threshold (smaller d_s) allows the accomodation of more transmissions, but increases the average interference, potentially causing more collisions. A smaller threshold (greater d_s) provides more protection but may limit the throughput. This effect is of particular importance near the sink, because its one-hop neighbors may be blocked even by far handshakes, if d_s is too high.

Figures 5.2 to 5.4 compare EPC^2 and CSMA/CA, showing the average throughput at the last hop, the average transmission success ratio, and the average end-to-end delivery delay, respectively. All metrics are drawn vs. the network packet generation rate, λ, for some values of d_s. We can observe that EPC^2 outperforms CSMA/CA in terms of throughput and delivery ratio for all values of the sensing range. In particular, more than 50% more packets per second are delivered to the sink, along with a better overall network success ratio, confirming EPC^2's ability to deliver traffic to the destination correctly. However, this efficiency comes at a price, as slotted transmissions require that nodes wait until their transmission turn arrives. Moreover, during EPC^2's transmit phase, one-hop neighbors of the sink cannot be contacted by other nodes. On average, this translates into a greater end-to-end delay, as seen from Figure 5.4. Notice that this delay is computed for the whole network, so that a double effect is shown: first the delay decreases, thank to the better delivery efficiency of EPC^2; then the amount of data to be delivered to the sink begins to increase with increasing λ, implying longer transmit and waiting periods during the transmit phase, and consequently an overall greater transmission delay. We wish to highlight that this is not necessarily an unwanted effect: it comes along with better throughput and delivery ratio, meaning that more traffic is more effectively delivered, only with a different organization of the transmissions. In the following comparisons, we will set $d_s = 15\,\text{m}$ as it is a good tradeoff between high throughput and a stable delivery ratio.

Tuning EPC^2: the length of the transmit phase

Following the reasoning developed up to this point, it is straightforward to understand the importance of the transmit phase length within an EPC^2 cycle. Recall from Section 5.2.1 that EPC^2 determines the number of packets to be received during the transmit phase according to the backlog length declared by each of its neighbors. As a rule of thumb, the longer the transmit phase, the nearer the throughput to the maximum channel capacity, thanks to the improved efficiency. On the other hand, a long phase also delays the delivery of data from more distant nodes.

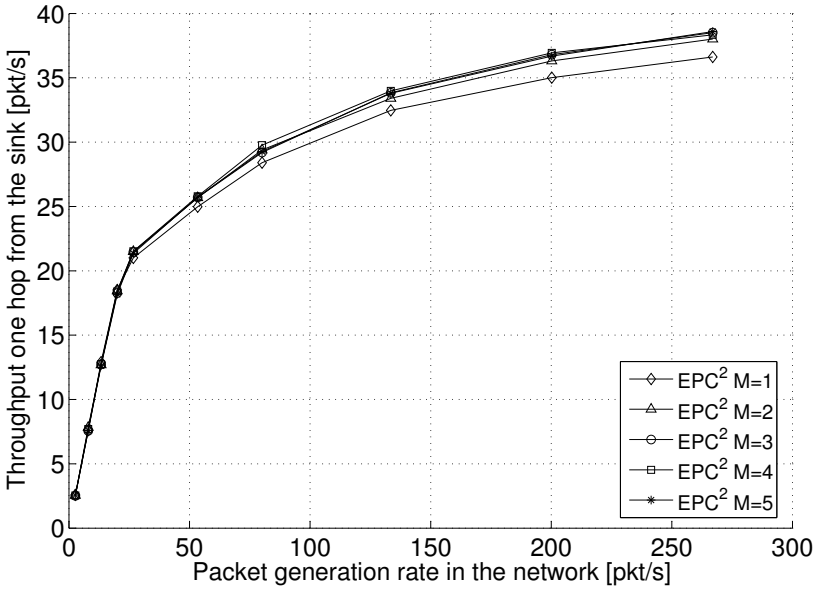


Figure 5.5: Throughput (last hop).

Recall that N_s is the number of one-hop neighbors of the sink, and M the parameter that specifies how many packets each of them is allowed to transmit during its turn. Therefore, the sink expects at most MN_s packet per transmit phase. The effect of varying M can be observed from Figures 5.5 to 5.7, which respectively depict the average throughput around the sink, the network transmission success ratio and the end-to-end packet delivery delay. In all pictures, M is increased from 1 to 5. From Figure 5.5, we can see how the throughput near the sink is not deeply affected by M. We may observe a slight increase for $M = 2$ with respect to $M = 1$, but increasing M further does not yield a substantial gain. The same considerations hold for the network success ratio (Figure 5.6), where the slightly better values achieved for $M \geq 2$ are due to the better throughput at the sink, which allows to deliver more packets and reduces the chance that some of them are dropped before reaching the sink. The end-to-end delay at the last hop depends a little more strongly on M, as the delay floor at high traffic increases from 12.2 to 13.1 s for increasing M. Based on the insight gained from these results, we decide to set $M = 2$ from now on, which seems again a good tradeoff that accomodates high throughput and low delay.

Performance under event-based traffic

To further test the robustness and delivery capabilities of EPC^2, we employ in the following the event-based traffic model described at the beginning of this Section. Namely, the network generates traffic according to a Poisson process with very low rate (0.1 pkt/s). During the simulation, some events are generated that are sensed by all nodes located within a certain

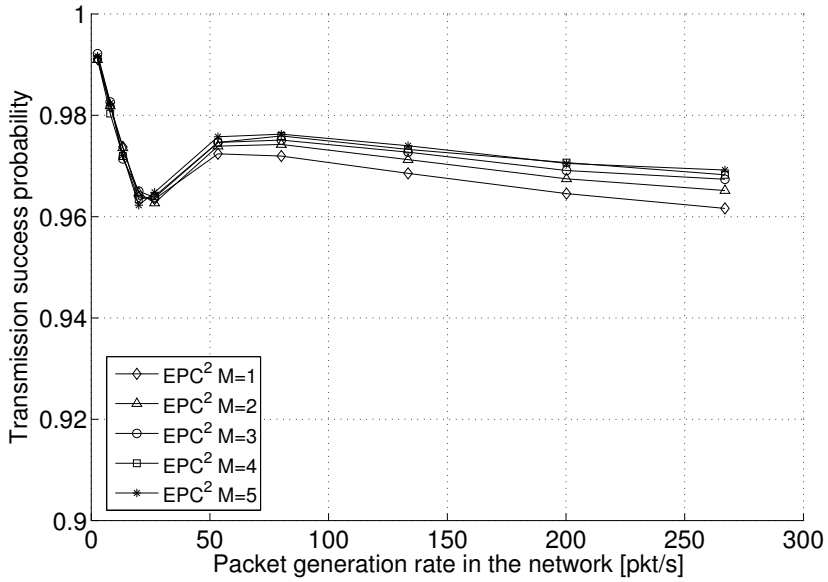


Figure 5.6: Transmission success ratio.

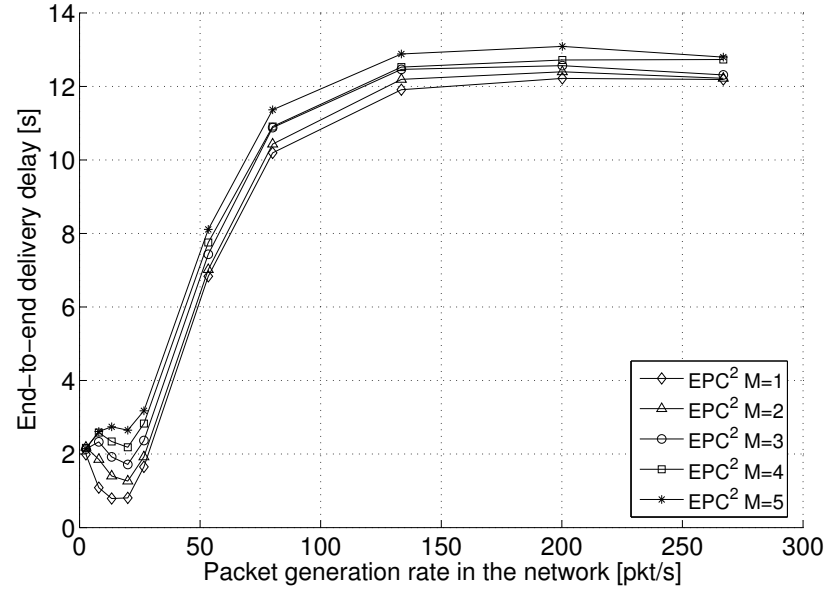


Figure 5.7: End-to-end delivery delay (last hop).

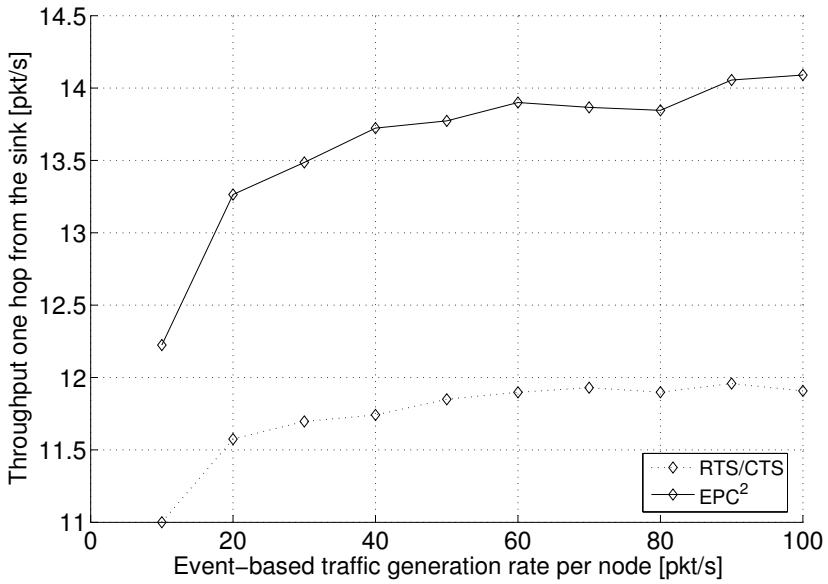


Figure 5.8: Throughput (last hop) under event-based traffic generation.

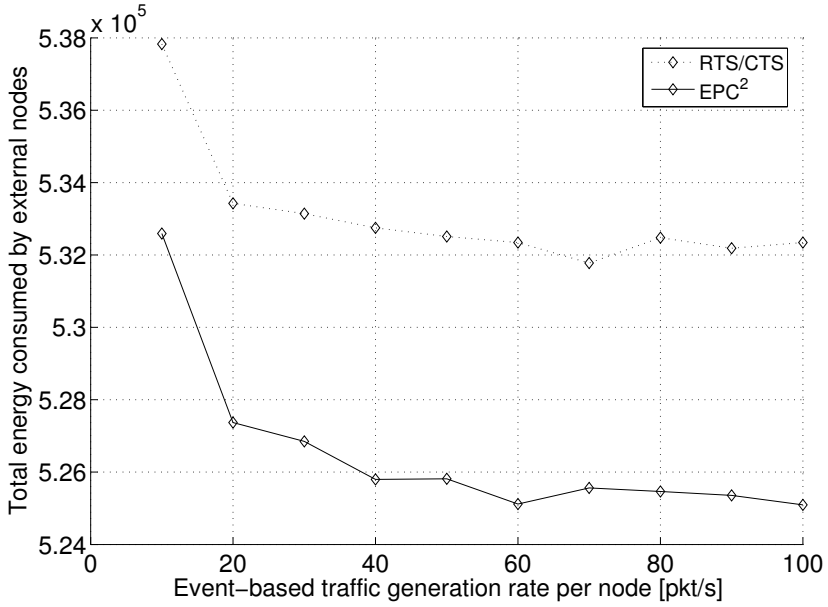


Figure 5.9: Normalized energy consumption under event-based traffic generation.

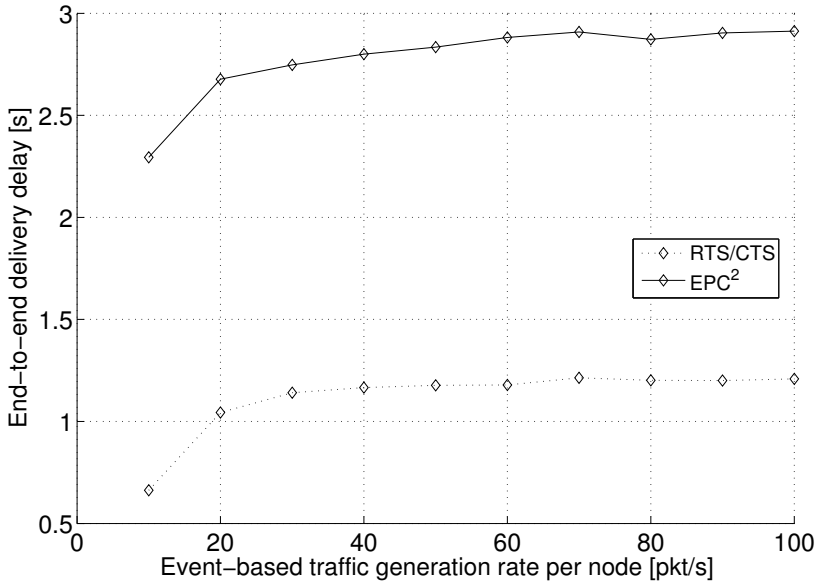


Figure 5.10: End-to-end delay under event-based traffic generation.

distance from the event (randomly chosen between 2 and 16 m). Such nodes generate data at a much higher rate, which we varied between 10 and 100 pkt/s per node. In particular, 100 pkt/s is the value chosen in [60]. As for the other results in this Section, we average the results over a number of network realizations to achieve statistically meaningful results.

Figures 5.8 and 5.9 respectively depict the throughput near the sink and the energy consumption per packet, normalized to the energy spent for transmitting one data packet, both vs. the packet generation rate upon event detection. It can be observed that the increased efficiency ensured by EPC^2 allows to deliver more packets to the sink even with bursty transmissions. Interestingly, the energy consumption is smaller for EPC^2 than for CSMA. The reason is twofold: on one hand, EPC^2 allows for an increased delivery efficiency near the sink, on the other hand, CSMA wastes energy for advancing packets that are dropped at some point due to a large number of failed retransmissions. In fact, nearly 27% of the packets are discarded for this reason by CSMA, against a mere 6% achieved by EPC^2. One of the disadvantages of our approach is that it requires the sink neighbors to wait nearly twice as long, on average, before their packets can be delivered (see Figure 5.10). This effect is due to the different organization of the transmissions, and seems a reasonable price to pay for the increased overall efficiency achieved.

5.2.3 Discussion

The results obtained from our study confirms that it is more convenient to interface EPC^2 with CSMA/CA rather than simply drive the whole network with the latter. In particular we

have shown that, allowing for the one-hop neighbors' transmissions to be scheduled (which tends to increase the delivery delay at the last hop), the throughput near the sink (and thus in the whole network) is increased along with a general transmission efficiency. Furthermore, even if our protocol requires some more effort, *e.g.*, to keep nodes synchronized, this additional complexity is confined to one-hop neighbors of the sink, which makes it more suitable for real-world implementation.

Chapter 6

Energy–efficiency in underwater sensor networks

Energy–efficiency of WSN is a key issue that permeates all the aspects of the network design, from hardware to protocols and applications.

Energy efficiency has importance both in radio and underwater networks, even though the parameters and the strategies adopted to pursue this objective can be substantially different due to the dissimiliarity in the transmission medium and hardware components. Schemes that provide energy efficiency can be implemented both at the physical layer, e.g., by choosing a particular modulation, and at the MAC and routing layers, by avoiding collisions and retransmissions or selecting paths that minimize the energy consumption.

In this chapter, we considered an underwater scenario. We tackled the energy efficiency problem both at the MAC layer and physical layer, and proposed two simple and distributed solutions, suited to the limitations of underwater sensor hardware. The main assumption of our scenario is the capability of nodes to adjust their transmission power in order to reach a relay or the destination node using as little power as possible.

First of all, we investigate the effect of duty–cycle in such a network, in order to understand whether this technique, widely used in radio networks to reduce energy consumption, also provides some benefit to underwater networks.

Second, we propose a very simple dynamic channel access scheme, that levers on the strong frequency dependency of the underwater channel response to improve energy efficiency.

6.1 State of the art

Although underwater sensor networks (UW) are a quite novel field of interest, a lot of research has been developed in recent years. One of the most widely addressed topic is energy efficiency. Several works investigated the behavior of radio protocols when applied to the

underwater scenario. In fact, the strong differences between the two scenarios, in terms of bandwidth, channel response, propagation delay and hardware constraints, may require new solutions whereas the schemes developed for the radio environment may be found inefficient.

Much efforts have been spent on MAC and routing protocol design. In [65] the authors aim at energy saving by introducing a collision avoidance scheme with RTS/CTS handshake. Knowing the distance between transmitter and receiver, nodes can set a waiting time for DATA transmission, according to the propagation delay, in order to be aware if other transmissions are ongoing. In this way it is possible to reduce collisions on data packets. In [66], a MAC algorithm based on sleep cycles is proposed. Every node sends a SYNC packet in order to syncronize its cycle. Therefore, knowing of the cycles of the other nodes, transmissions can be scheduled in an optimal manner. However, this algorithm requires the network to be static in order to keep nodes' syncronization.

Underwater routing problems are addressed in [67], that takes into account the problem of the relationship between distance and energy transmission, as in [68] and [69]. A geographic routing protocol called Focus Beam Routing is proposed in [70]. The purpose is to shorten hops in order to save energy using different transmission power levels.

In [71], duty–cycle and wake–up radio are compared, but nodes are assumed to have a fixed transmission power.

6.2 Underwater channel model

Underwater communications are usually based on acoustic waves because electromagnetic propagation suffers a severe attenuation resulting in a very small coverage range, not suitable for a sensor network scenario. Several factors affect the underwater channel response, like waves, depth, temperature and ships. Therefore it is quite difficult to include every single contribution in a model that is, in most cases, roughly approximate.

In this section we describe the main features of an underwater channel, defining the model that we used in our study.

The attenuation experienced by a signal at frequency f over a distance d follows the equation (in linear scale)

$$A(d, f) = A_0 d^k a(f)^d, \tag{6.1}$$

where A_0 is a unit-normalizing constant, k is the spreading factor (whose practical value is tipically taken to be 1.5) and $a(f)$ is the absorption coefficint.

The value of this coefficient is given empirically by Thorp's formula [72],

$$10 \log a(f) = 0.11 \frac{f^2}{1+f^2} + 44 \frac{f^2}{4100+f^2} + 2.75 \cdot 10^{-4} f^2 + 0.003 \tag{6.2}$$

where the coefficient is expressed in dB/km and the frequency f is in kHz.

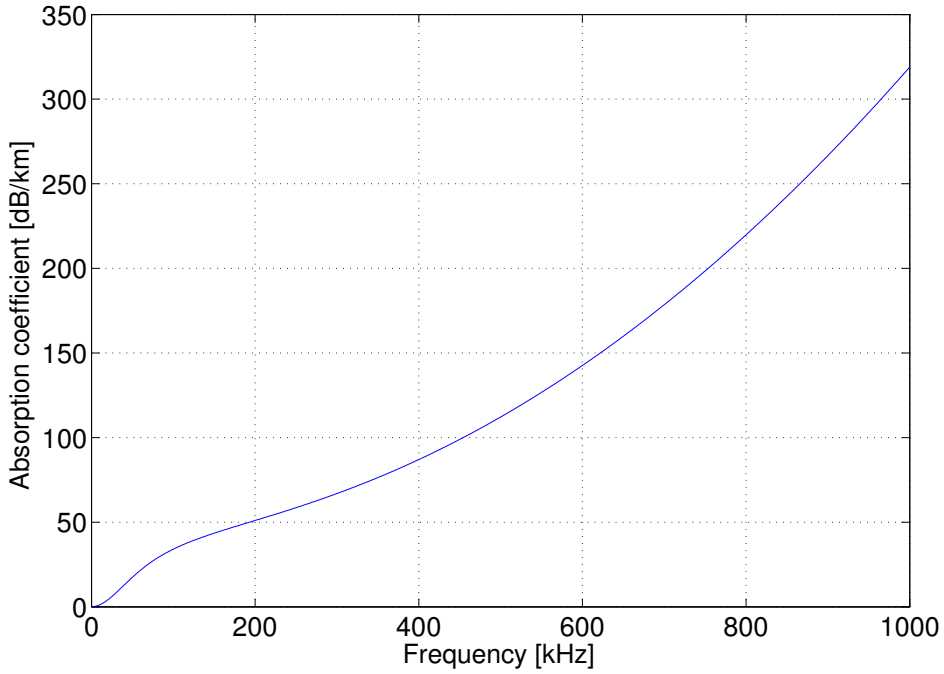


Figure 6.1: Absorption coefficient

As shown in Fig. 6.1, the absorption coefficient grows very quickly as the frequency increases. This is the main limiting factor to the available bandwidth in underwater communications. Therefore, the carrier frequencies are usually set around $5 \div 50$ kHz even though the actual limit strongly depends on the distance at which the transmission occurs.

The noise experienced on an underwater acoustic link is composed of four main sources: turbulence, shipping, waves and thermal noise. Each of them has a particular expression in dB re μPa per Hz [73]:

$$
\begin{aligned}
10 \log N_t(f) &= 17 - 30 \log f; \\
10 \log N_s(f) &= 40 + 20(s - 0.5) + 26 \log f - 60 \log(f + 0.03); \\
10 \log N_w(f) &= 50 + 7.5\sqrt{w} + 20 \log f - 40 \log(f + 0.4); \\
10 \log N_{th}(f) &= -15 + 20 \log f;
\end{aligned}
$$

where s is the shipping factor (between 0 and 1, for low or high activity respectively), w is the wind speed in m/s and f is the frequency in kHz.

Moreover, these components have different impact at different frequencies: the turbulence noise is interesting only in a very low frequency region, less than 10 Hz, while shipping is dominant around 10-100 Hz. The wave noise has a strong impact in the bandwidth 100 Hz - 100 kHz, while the thermal noise only for $f > 100$ kHz.

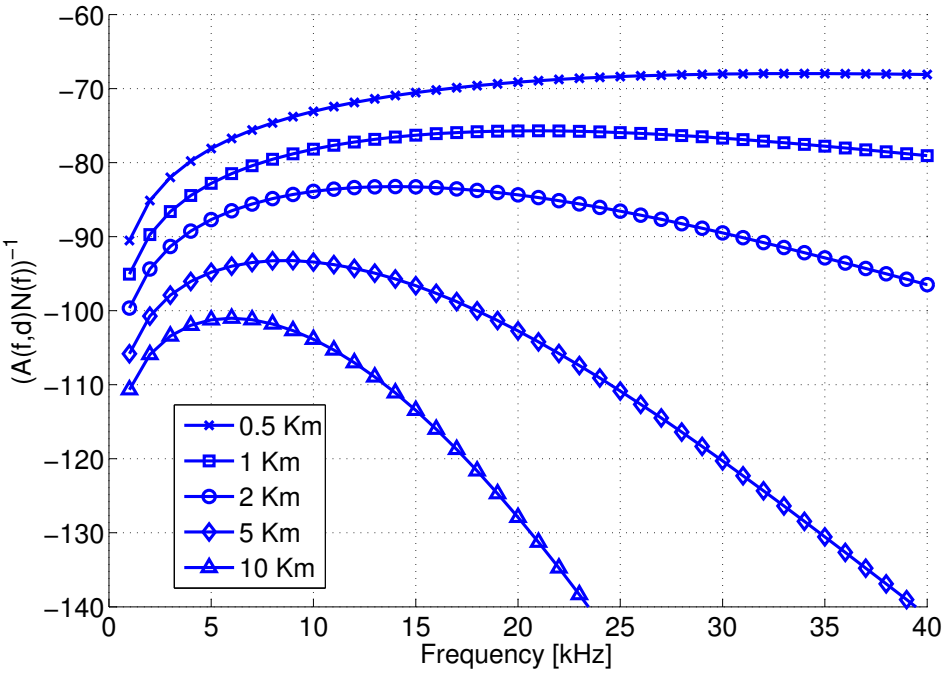


Figure 6.2: $\frac{1}{AN}$ for different frequencies and distances.

Focusing on typical used bandwidth and considering an average behaviour, it is possible to find an approximation of the overall environmental noise in order not to deal with all these different parameters. In our study we use the approximation expressed in [69],

$$10 \log N(f) \approx N_1 - \eta \log f \tag{6.3}$$

where $N_1 = 50$ dB re μPa and $\eta = 18$ dB/decade.

It is worth to notice that both attenuation and noise are strongly dependant on the frequency. Moreover the attenuation is affected by the distance both through the spreading factor exponent (similar to the path loss exponent in radio channel) and as the exponent of the absorption coefficient, making the path loss more severe compared to that found in the radio environment.

Fig. 6.2 shows the value of $\frac{1}{A(f)N(f)}$, which is proportional to the SNR, for different frequencies and distances. The behavior of the SNR differs from that of radio communications, and strongly characterizes the underwater communications.

Another peculiarity of acoustic channel is the time scale. The propagation speed of sound in the water is very slow. Acoustic waves propagates in water with a speed $c_s \simeq 1500$ m/s, much slower than the speed of electromagnetic waves in a radio environment. Moreover the bandwidth is limited to few kHz, so that only transmission rate in the order of kbit/s can be supported.

All these characteristics of the underwater channel call for a different approach compared to radio sensor networks and pose new challenges in the design of protocols and techniques for the underwater scenario.

Therefore, it is necessary to study and, if necessary, adapt the energy–efficiency algorithms used in radio wireless networks to the underwater scenario and find new approaches in order to save energy in the network.

6.3 Simulation scenario

We consider a network of N nodes deployed over an area A. In a conventional setting, these nodes would transmit packets when the channel is available, i.e., according to the MAC and routing protocols used. In the proposed setting, a duty cucle, $\alpha \leq 1$, is introduced, such that each node is "on" for the fraction α of time, and "off" for the remaining time. When it is off, a node can neither transmit, receive, nor listen. Specifically, we use a fixed interval of time T_c for one cycle. This interval is the same for all the nodes, but each node chooses its active time αT_c within the cycle, randomly and independently of the other nodes. The duration of a cycle is chosen longer than the packet length, and a node already engaged in transmission is not interrupted if the cycle runs out before the transmission is complete.

Network traffic is modeled as a Poisson process with rate $1/\tau$, where τ is the mean inter-arrival time between two consecutive packets. Consequently, every node generates packets following a Poisson process with rate $1/(N\tau)$, independently of one another. Traffic is addressed to a common sink, placed in the center of the network.

The nodes are deployed, as shown in Fig. 6.3, according to a random grid method as follows: the area is divided into a grid of cells, and a node is placed in each cell according to the uniform distribution. Therefore, the inter–node distance is bounded by $d\sqrt{5}$, where $d = \sqrt{A/N}$ is the side of a cell. The network density is $\rho = \frac{N}{A}$.

When the duty cycle α is used, the effective number of nodes, i.e. the average number of active nodes is $N_{eff} = \alpha N$, and, hence, the effective density is

$$\rho_{eff} = \frac{N_{eff}}{A} = \alpha \cdot \rho. \tag{6.4}$$

Every node has a finite number L of power levels ($L = 4$ in our scenario). Their values are set in accordance with the effective node density, i.e., in accordance with the total number of nodes N and the duty cycle α. The power levels are chosen to correspond to nominal transmission distances

$$D_L = [d_{min} : d_{step} : d_{max}], \tag{6.5}$$

where $d_{min} = d$, the side of a cell, and

$$d_{max} = \min(\sqrt{5\rho_{eff}^2}, \sqrt{A} \cdot \frac{\sqrt{2}}{2}). \tag{6.6}$$

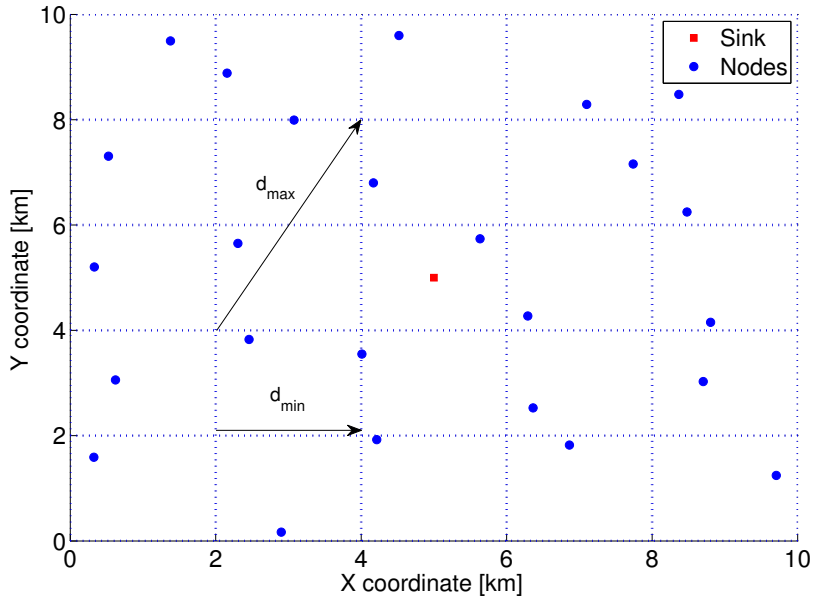


Figure 6.3: Network deployment for 25 nodes.

It is worth noting that this power allocation provides a fully connected network in case of $\alpha = 1$. Using power levels equally spaced in distance, as proposed in [70], we have that

$$d_{step} = \frac{d_{max} - d_{min}}{L - 1}. \tag{6.7}$$

Given a transmission bandwidth B, every distance is hence associated to a particular transmission power according to the attenuation law and target SNR, as follows,

$$P_i(\text{dB}) = SNR + A(D_i, f, B) + N(f, B) + B \tag{6.8}$$

where $A(D_i, f, B)$ and $N(f, B)$

$$A(D_i, f, B) = \frac{1}{B}\int_B A(D_i, \nu)d\nu \tag{6.9}$$

$$N(f, B) = \frac{1}{B}\int_B N(\nu)d\nu \tag{6.10}$$

are the average of (6.1) and (6.3) over the bandwidth B.

To evaluate the energy consumption, it is necessary to convert the pressure value P_i into Watt. We consider that a reference acoustic pressure P_{ref} of 172 dB re μPa at 1 meter is produced by 1 W. Therefore it is simple to describe the conversion to Watt as

$$10 \log P_i(W) = P_i(\text{dB}) - P_{ref} \tag{6.11}$$

6.4 Effect of duty–cycle in UW

Nodes density has a strong impact on network performance. Energy consumption, congestion, delay and collisions are strictly related to the topology of the network. In our study, we focused on the energy consumption aspect. Thus we investigate how the network topology affects the energy performance. We also consider other performance metrics in order to have a more general view of the network behavior, identifying trade–offs between energy saving and reliability and latency.

In our simulations, we studied many scenarios for different values of ρ_{eff}, . Hence, we varied both the numbers of nodes and different values of the duty–cycle parameter.

The duty–cycle affects the network in different ways. First of all, it reduces the ρ_{eff}, because this metric takes into account only the active nodes, leading to longer hops on average. Moreover, it allows to energy saving, because during the sleep time the energy consumption of a node is negligible. Also, the management of the network is more complicated. The randomized selection of the active nodes, based on the duty–cycle, has the same effect of a variable topology.

6.4.1 Results

Results are obtained using the AUVNetSim Python Network Simulator [74].

In our simulations, the network supports three different MAC protocols: CS-ALOHA, DACAP [65] and CSMA-CA protocol. As a routing protocol, we use FBR [70], that supports different power levels and is suitable for our dynamic topology scenario.

Energy Performance

As previously mentioned, ρ_{eff} affects the energy consumed for transmissions ε_{tx}, receptions ε_{rx} and idle–listening ε_{id}. Whereas the comparison among the energy consumption of a specific operational mode (transmission, reception, idle–listening) in different scenarios is not related to the power value, the study of the whole energy consumption in the network ε_{tot} is more complicated and a very important role is devoted to the power values.

Moreover, even the energetic efficiency is a very important metric, calculated as the ratio between energy consumption per delivered packet and the number of nodes as follows

$$\gamma = \frac{\varepsilon_{tot}}{N} \tag{6.12}$$

This metric is inversely proportional to the lifetime of the network.

Therefore, we first analyze the energy consumption separately for transmission and reception. Second, we consider the overall energy consumption performance.

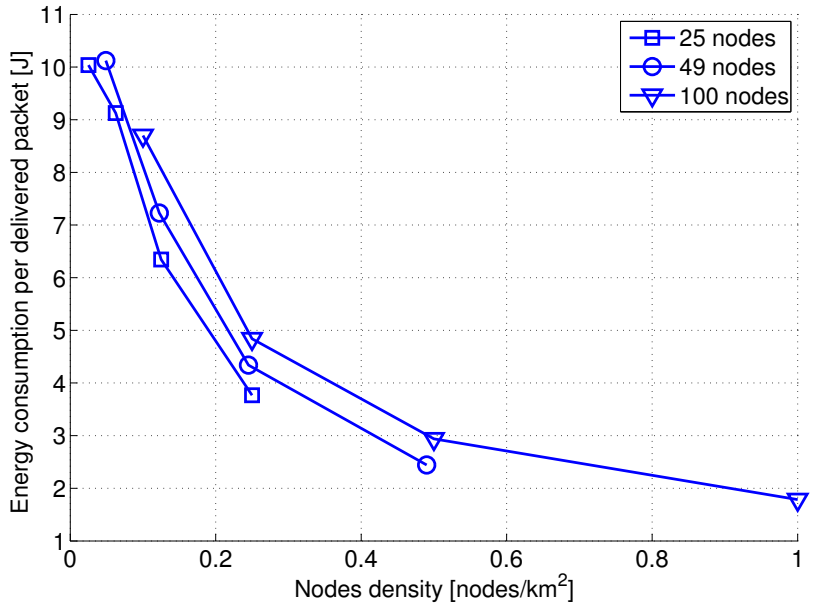


Figure 6.4: ε_{tx} as a function of ρ_{eff}

In the following graphs, we will show the results obtained with one MAC protocol only, for the others perform similarly, both in terms of absolute value and general trend. Thus, for the sake of clarity, we plot in the graphs only the results for CS–ALOHA MAC protocol.

In Fig. 6.4, we can see the transmission energy per delivered packet ε_{tx} as a function of ρ_{eff}. As expected, increasing ρ_{eff}, the transmission energy consumption decreases. Although shorter links mean a larger number of hops, we see that the relationship between distance and transmission power makes more advantageous to use short hops, leading to a comprehensive energy gain. If ρ_{eff} is the same, then networks with more nodes, that corresponds to a lower α, have worse energy performance. In fact, a low α means that the number of active nodes of the network changes over time. Therefore, it is more likely that some links break, requiring the routing protocol to run the discovery process again, thus wasting energy and time.

Focusing on the reception power, the relationship between density and energy consumption is reversed. At the receiver, the energy consumption ε_{rx} is not affected by the distance. Hence increasing ρ_{eff} leads to a greater number of transmissions with an increase of ε_{rx}. Regarding to the effect of α, the results confirms what we observed previously: a low α causes routing inefficiencies.

It is also interesting to look at the idle–listening energy consumption ε_{id} per delivered packet. Even here we can decouple the effects due to the number of nodes and the duty–cycle. In the first case, adding nodes obviously leads to an increase of the energy consumption from a network perspective. Therefore, if the number of nodes doubles, the network doubles the

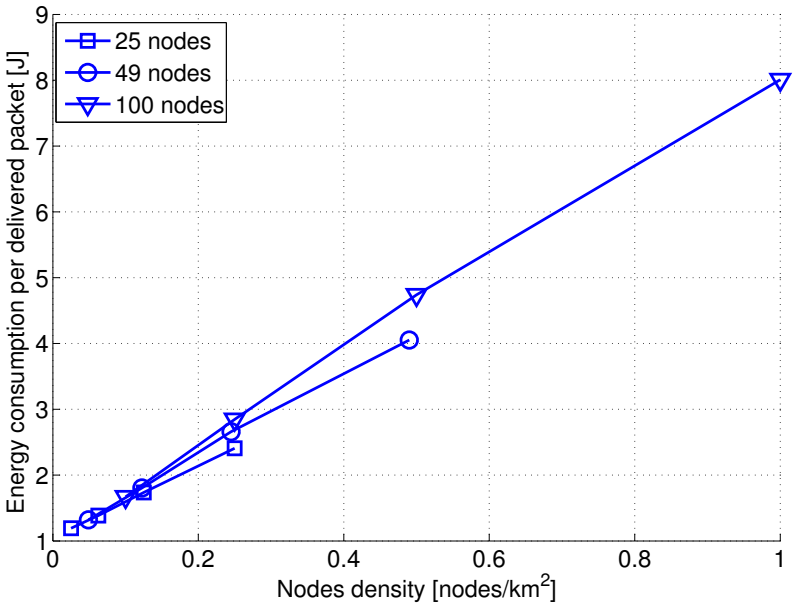


Figure 6.5: ε_{rx} as a function of ρ_{eff}

	Ref TX power	RX power	IDLE power
Low efficiency transmission	152 dB re μP	1 W	1mW
High efficiency transmission	172 dB re μP	1 W	1mW

Table 6.1: Power values for the different power scenarios

available energy, but also the energy consumption for idle–listening doubles. The effect of duty–cycle, instead is the opposite: if $\alpha = 0.5$, then ρ_{eff} halves and ε_{id} also halves, while the available network energy remains the same. Therefore, ε_{id} depends linearly only on ρ_{eff}.

At this point it is interesting to understand how the total energy consumption ε_{tot} is affected by the number of nodes and the duty–cycle. This analysis strongly depends on the amount of power needed for each operation: transmission, reception, idle-listening.

We analyze here two scenarios with different relationship among the power values. As previously stated, we consider a reference transmission power P_{ref} of 1 W, which produces an acoustic pressure of 172 dB re μPa. In addition to this power, we consider a *loss* value that takes into account some inefficiencies in the device components and extra attenuations, that force the transmitter to increase the power. We can obtain the same effect decreasing the reference power as in Tab. 6.4.1, where the two different sets of power values are presented. Obviously, the higher P_{ref}, the less power is required to transmit.

In Fig. 6.6, we can notice that if the device has not very efficient transmission hardware,ε_{tot} is minimized around 0.25 nodes/km^2. For very low values of ρ_{eff}, the transmission energy

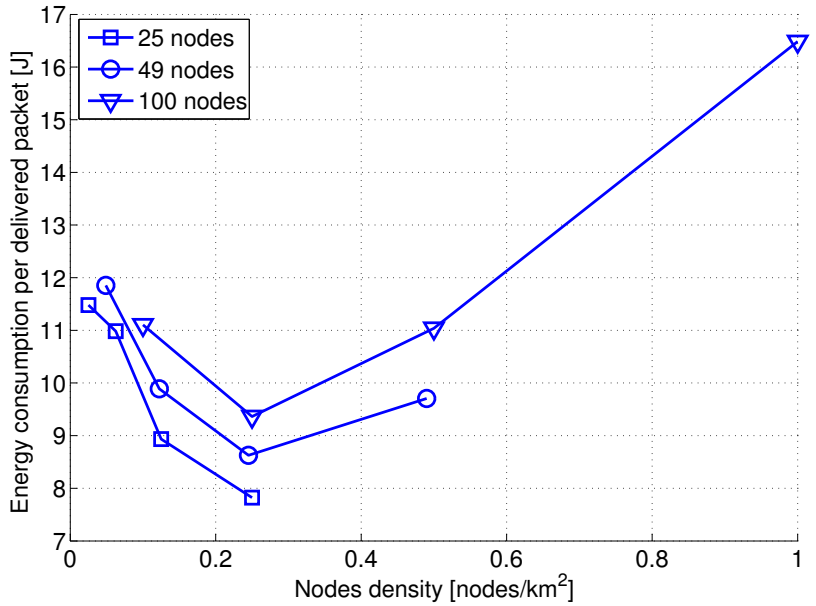


Figure 6.6: Total energy consumption with high transmission power

saving due to higher nodes density is more significant than the increase of ε_{rx} and ε_{id}. For densities larger than 0.25, even if ε_{tx} decreases, the gain is not enough to compensate the increase of ε_{rx} and ε_{id}.

If ε_{tx} is very efficient, it means that the main source of energy consumption is the receving and idle–listening one. Therefore, ε_{tot} increases with ρ_{eff} in an almost linear way, due to the quasi-linear behaviour both of ε_{rx} and ε_{id}.

Furthermore, we analyzed the efficiency metric described in (6.12).

The two graphs Fig. 6.7 and Fig. 6.8 show γ in the two power scenarios of Tab. 6.4.1.

In Fig. 6.7, the number of nodes is the main parameter that affects the energy consumption, because it strongly affects the distance among nodes and the possibility to use lower power levels. Therefore, regardless of the duty–cycle value, γ is increased by increasing the number of nodes. Then, given the number of nodes, it is possible to choose the optimal value of α, that in our case, is such that $\rho_{eff} = 0.25$ nodes/km^2.

For an efficient transmission device, the number of nodes becomes as important as the duty–cycle. In fact, the duty–cycle enables to save a lot of energy during the idle listening status that, in this scenario, is the dominant source of energy consumption. For example, fixing α and doubling the number of nodes, ε_{tot}/N slightly decreases, whereas decreasing α yields much larger power saving. The relationship between ε_{tot} and α is almost linear, with a slope that increases when the number of nodes decreases.

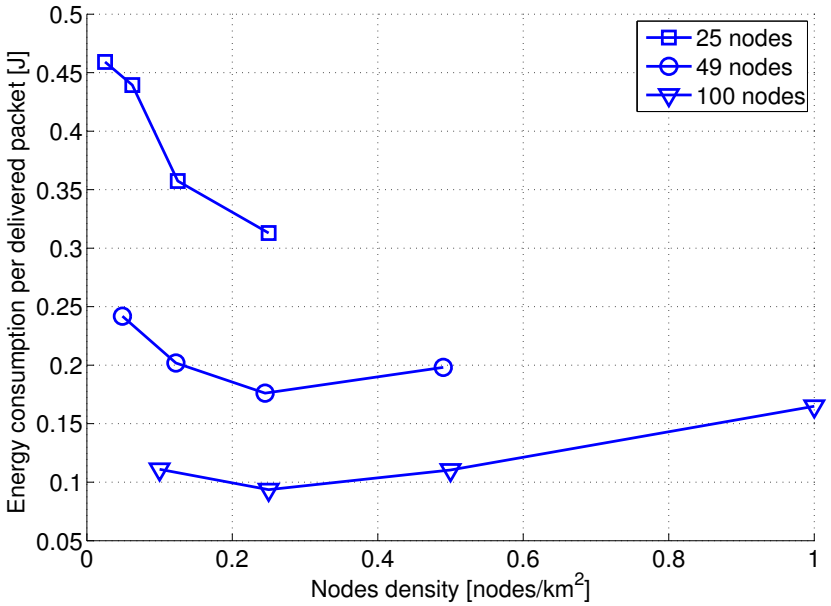


Figure 6.7: ε_{tot}/N as a function of ρ_{eff} in high power transmission scenario

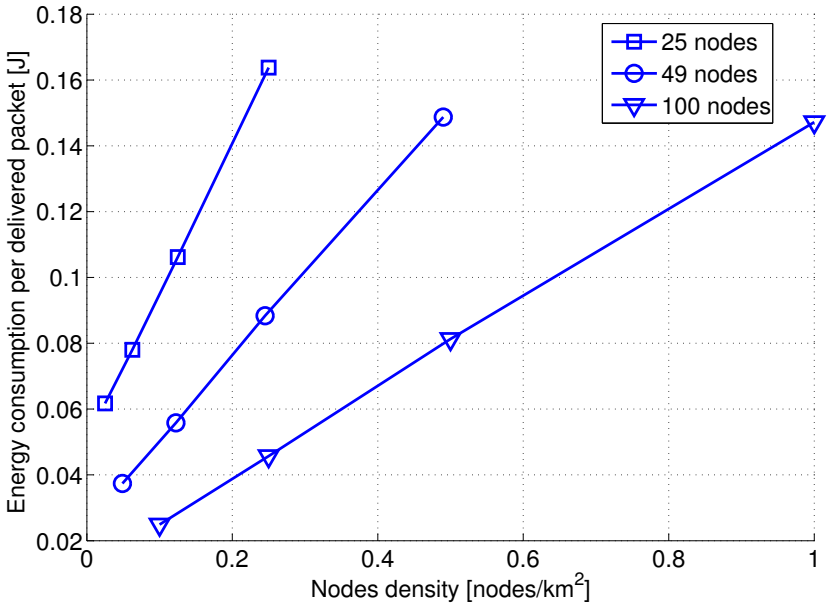


Figure 6.8: ε_{tot}/N as a function of ρ_{eff} in low power transmission scenario

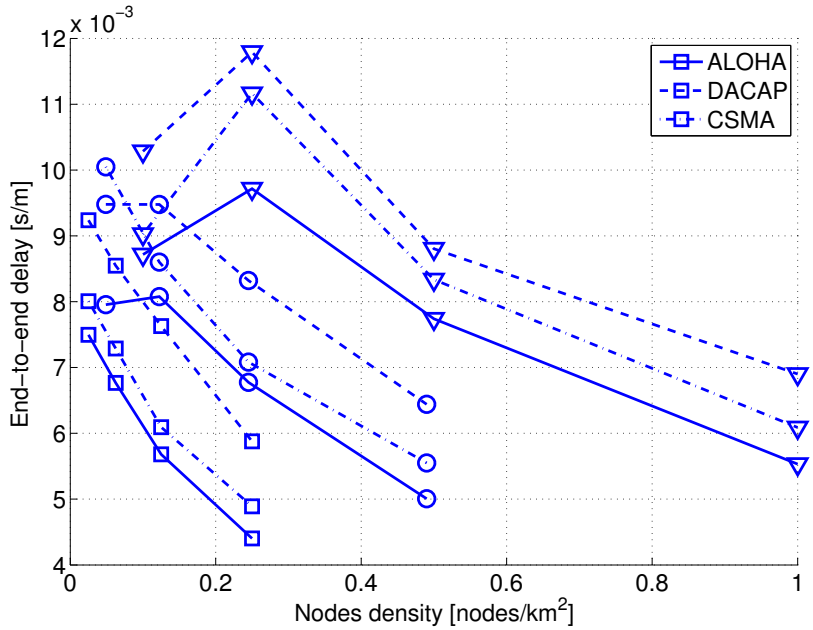


Figure 6.9: Packets delay per meter - The triangles represent the scenario with 100 nodes, the circles with 49 nodes and the squares with 25 nodes

Network Performance

As a second step, we analyze other perfromance metrics, as delay, packet delivery ratio and collisions. Even in this case, we investigate if whether or not these performance indexes are affected by the number of nodes and the duty–cycle.

Fig. 6.9 shows the delay performance for different MAC protocols and different values of ρ_{eff}. The graph clearly points out that both the number of nodes and the duty–cycle affect the end-to-end delay.

Adding nodes increases the number of hops and the total delay, while a low duty–cycle increases the probability of broken paths, adding delay.

Concerning the MAC protocols impact, we aobserve that, as expected, ALOHA is the fastest protocol to deliver packets even though the usage of FBR routing protocol adds some redundancy, due to the RTS/CTS handshake used for route discovery. On the contrary, DACAP, with its waiting times to avoid collisions, takes more time compared to the other MAC protocols.

We are also interested in the reliability of the network. Therefore we analyzed the packet delivery ratio, i.e. the ratio between the number of packets arrived to the sink and the number of packets generated by the network:

$$\xi = \frac{\text{Number of delivered packets}}{\text{Number of generated packets}} \tag{6.13}$$

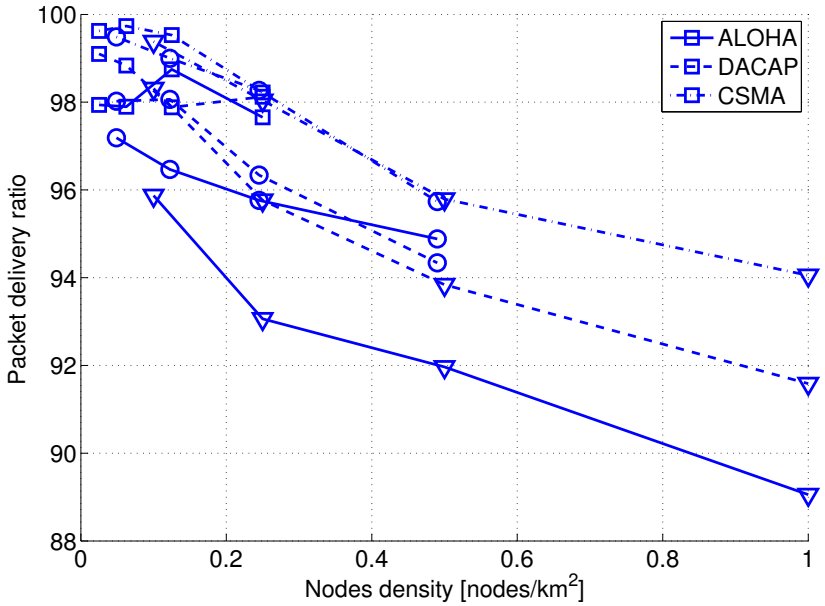


Figure 6.10: Packet delivery ratio - The triangles represent the scenario with 100 nodes, the circles with 49 nodes and the squares with 25 nodes

From Fig. 6.10 we see that ξ decreases as ρ_{eff} increases. When ρ_{eff} grows, the network performs more transmissions, causing an increment of packet collisions, that leads to a larger number of dropped packets because of retransmissions. In fact, in our simulator a packet can be retransmitted only a limited number of times, set to $N_{retx} = 4$ in our simulations. We note that this problem depends on the MAC.

CSMA and DACAP, which use a handshake before transmitting, are quite reliable: even in a very dense network they yield $\xi \geq 0.9$ and for low density $\xi \geq 0.96$. The delivery performance is affected only by ρ_{eff}.

Instead using ALOHA, which does not provide handshake (in our scenario handshake is used by FBR during the route discovery process), the sink receives from the network less than 90% of the packets in the worst case. Moreover, it behaves differently in scenarios with different number of nodes. Therefore, both a large number of nodes and a low duty–cycle decreases this reliability performance.

Hence, a low density network is the best scenario to have good delivery ratio. In the case of ALOHA, also a high duty–cycle is needed to improve the reliability.

The last metric we analyzed is the number of data packet collisions. It is important because energy consumption, delay and reliability are all affected by this metric. In fact, more collisions usually mean more energy consumption because of the higher number of transmissions, more delay and more dropped packets because of the retransmissions, due to

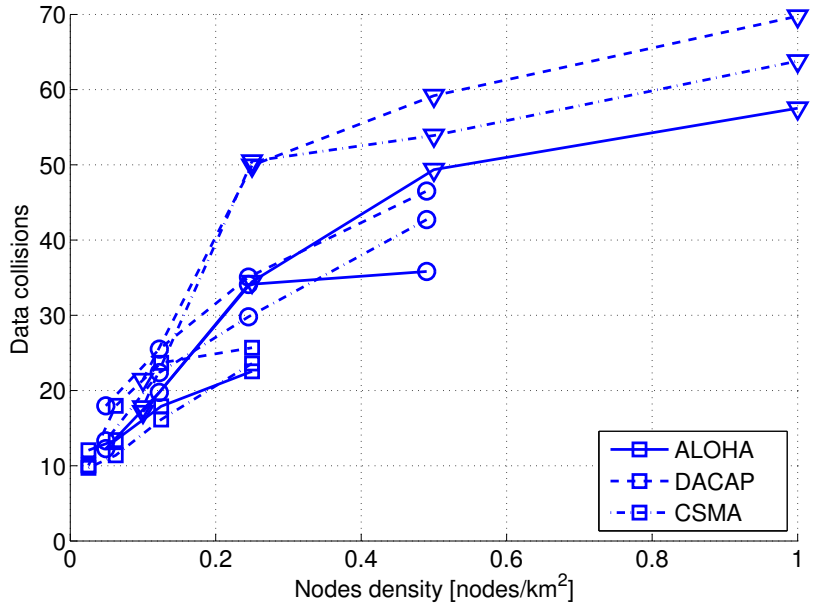


Figure 6.11: Data packet collisions - The triangles represent the scenario with 100 nodes, the circles with 49 nodes and the squares with 25 nodes

the limit on the number of retransmissions for a single packet as stated before.

As expected, Fig. 6.11 shows that collisions increase with increase ρ_{eff}. Unexpectedly, ALOHA yields the least number of collisions, while DACAP has the worst performance. It is also interesting to notice that in this graph duty–cycle affects the performance: high duty–cycle leads to less collisions.

From this plot and the previous ones, we can conclude that the collisions in DACAP and CSMA are not harmful in the sense of reliability, because the delivery ratio has good performance, but only in terms of energy and delay. Therefore DACAP and CSMA are more robust to collisions: if a collision happens, they can recover and transmit correctly the next attempt. On the contrary, ALOHA does not have this capability, thus each collision, yields with some probability packet dropping. This behavior gives some advantage to the delay performance that is calculated only on the packets received by the sink.

6.4.2 Discussion

In conclusion, we have observed that transmission energy consumption is reduced in dense networks, while low values of nodes' density provides better performance in terms of reception and idle–listening energy efficiency.

Instead aggregate energy efficiency is strongly dependent on the power values given to the transmission, reception and idle–listening operations. If the transmission is very efficient,

i.e. low power is needed to reach long distances, it is better to have long hops to decrease the number of receptions thus preferring sparse networks. Asymptotically, if it is possible to reach every node with transmission power equals to 0, the best density tends to 0. Increasing the transmission power required to generate the pressure wave, it is possible to find a minimum that shifts into higher density as the transmission power increases, reaching an asymptotical value of infinite density.

From the energy point of view, all the studied MAC protocols perform almost in the same way.

From the network point of view, the increase of ρ_{eff} worsens the performance. As expected ALOHA works better in terms of delay but the reliability is not as good as in CSMA and DACAP.

Therefore, it is very important to accurately design the network parameters, in terms of number of nodes and sleep cycle, based on the area and the power characteristics of the devices, to have better energy efficiency and longer network lifetime.

6.5 Efficient bandwidth allocation

Assuming that the channel is static, given a certain bandwidth, it is possible to find for each power level a maximum reachable distance after that it is necessary to increase the power. Leveraging the frequency dependance of the channel attenuation, we want to study if it is possible to reach a longer distance with the same power (or analogously the same distance with a lower power) only by switching to a different part of the spectrum. The scheme is based on the following principles:

- Given a bandwidth B and a center frequency f_c, the nodes can use the entire bandwidth, the lower part $B_0 = [f_c - B/2\ ;\ f_c]$ or the upper part $B_1 = [f_c\ ;\ f_c + B/2]$. This is the simplest case with two available sub-bands and can be easily generalized.

- Nodes can exchange bandwidth information using the control packets used by the FBR routing protocol.

- When nodes use a sub-band, the bit rate is proportionally reduced. The noise power is also reduced.

- Sub-band transmission of a packet is chosen only if the total transmission energy for that packet is less than when using the entire bandwidth.

- If the minimum transmission energy is the same in both sub-bands, the channel that gives the best SNR is chosen (Note that both SNR are above threshold, but they are not the same due to the finite set of power levels available).

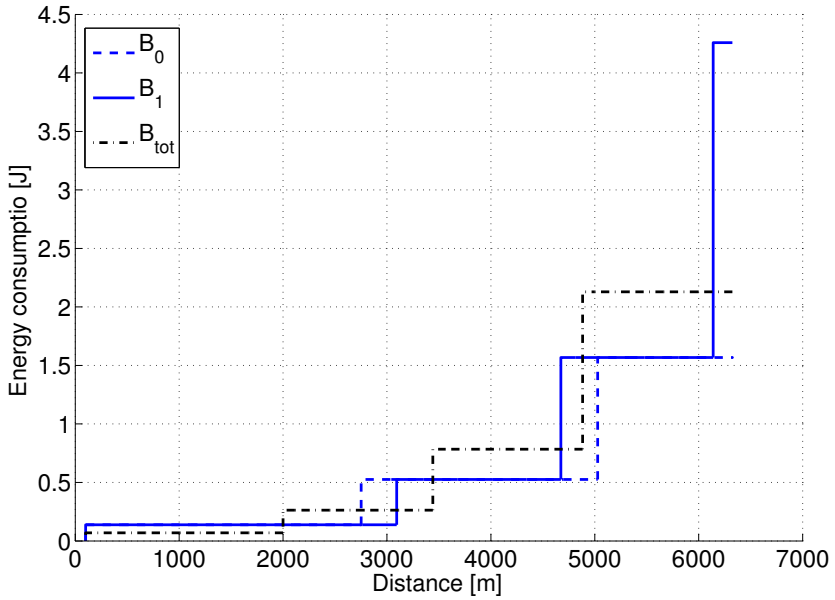


Figure 6.12: Low density scenario (N=25, α=0.5)

- If among the relay candidates, obtained from the FBR protocol, more than one node allows minimum power to be used, the nearest to the sink is chosen.

The advantage derived by the use of a sub-band depends on the distribution of the power levels P_L that actually depends on the density of the network. When a sub–band is used, the bitrate halves and the packet duration doubles, thus the energy consumption is $2P_i$, where P_i is the i-th power level. Hence, in order to have some gain using sub–bands, it is necessary that the channel response allows for transmit at a lower transmission power level P_{i-1} such that $P_i \geq 2P_{i-1}$ is verified.

From this policy, it follows that whenever possible, it is more convenient to tranmsit at the minimum power over the whole bandwidth.

The energy consumption performance using the proposed policy in two scenarios is reported in Fig. 6.12 and Fig. 6.13, for the case of low and high density scenario, respectively. Given a certain distance, the lower line is the best spectrum choice in terms of transmission energy. It is worth to notice that often the efficiency of B_0 and B_1 is the same. In Fig. 6.12, where density is low, this scheme can achieve better performance than in Fig. 6.13 for two reasons: first, the power levels are well spaced, therefore the advantage to use a lower power level is higher, then increasing the distance between nodes the channel becomes more variable with the frequency and the two sub–bands can have different attenuation responses.

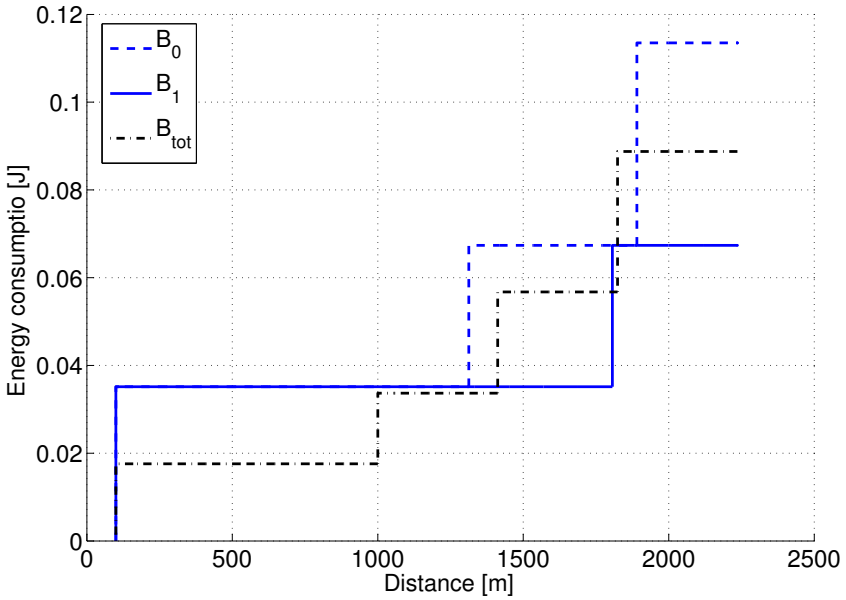


Figure 6.13: High density scenario (N=100, α=1)

6.5.1 Results

The presented results are obtained using AUVNETSim Python simulator [74] with CSMA MAC protocol and FBR routing protocol.

A quick look at DACAP shows that the performance is very similar to that of CSMA. Instead ALOHA provides slightly worse energy efficiency.

We analyze the proposed scheme, investigating the energy consumption, both for transmission and reception side.

As a first step, we analyzed separately the transmission and the reception energy consumption. In fact, the behavior of these metrics is independent of the particular power value used in the simulations. On the contrary, their impact in the total energy consumption strongly depends on the values that are used for modeling the devices.

Fig. 6.14, Fig. 6.15 and Fig. 6.16 show the difference on the transmission energy consumption per delivered packet for the two different bandwidth allocation schemes, the static one (Single channel) and the variable one (Double channel). The x-axis represents ρ_{eff}, that is the effective density of the network.

The gain that it is possible to achieve is between nearly 0% in the case of 100 nodes with duty–cycle $\alpha = 1$ and 40% in the cases with low density. This result is consistent with the theoretical improvements presented in the previous section, where it was seen that the best performance can be obtained in low density scenarios.

Indeed, in this case, the power levels are largely spaced so that a lower power level yields

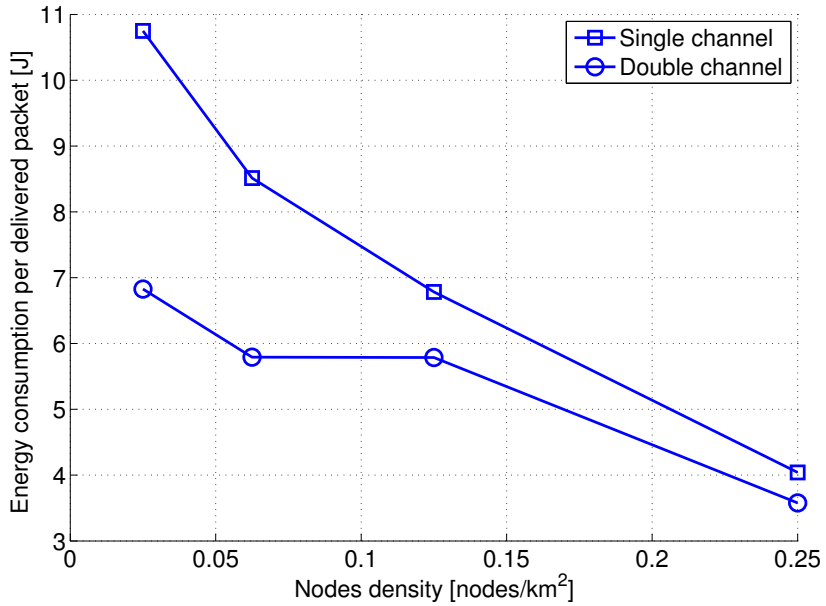


Figure 6.14: Transmission energy consumption per delivered packet with 25 nodes

effective energy saving even though the transmission time is twice longer. Moreover, in such a scenario, power levels are also chosen to reach long distances and, as shown in Fig. 6.2, at these distances the channel is much more frequency selective. Therefore it is necessary to dynamically choose the transmission bandwidth and the achievable gain can be very high.

Fig. 6.17 shows the difference on the receiving energy consumption for the two different bandwidth allocation scheme, as a function of ρ_{eff}, in the scenario with 49 nodes. However, even changing the number of nodes, the behavior of the receiving energy conmsumption is almost the same.

Keeping into consideration the receiving power, it is possible to see that the adaptive scheme increases energy consumption. Hence with the new scheme some transmissions have a longer duration and the reception power does not change, so the energy consumption increases proportionally to the packet duration.

The energy consumption in the dynamic scheme increases of a factor that varies between 1.5 and 2.

The total energy consumption is reported in Fig. 6.18. It is important to notice that there is a point at which the performance of the two schemes intersect, so that in certain conditions the static one is better than the adaptive and in other conditions the dynamic one achieves better results. As a general law, for low densities, in order to save energy, it is better to improve the transmission energy consumption, using the adaptive scheme, while for high densities it is better to optimize the receiving energy consumption, adopting the static

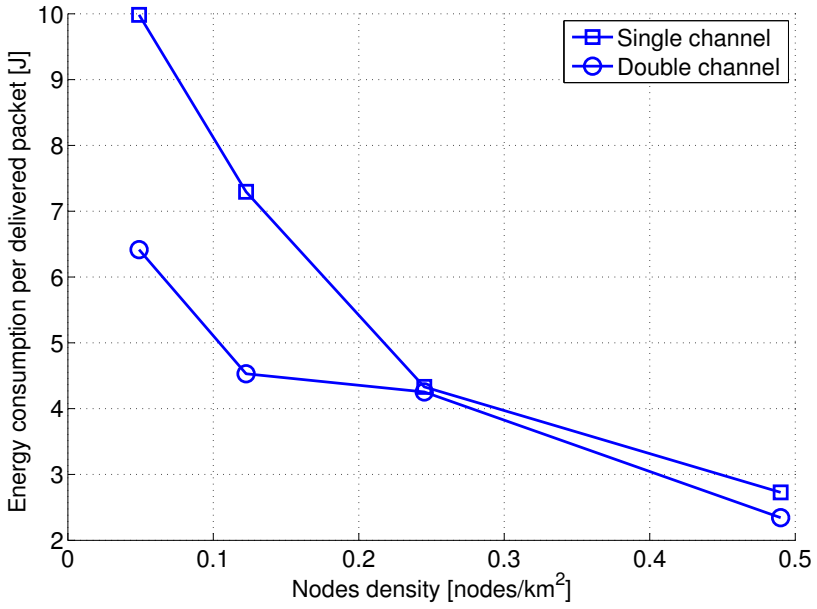


Figure 6.15: Transmission energy consumption per delivered packet with 49 nodes

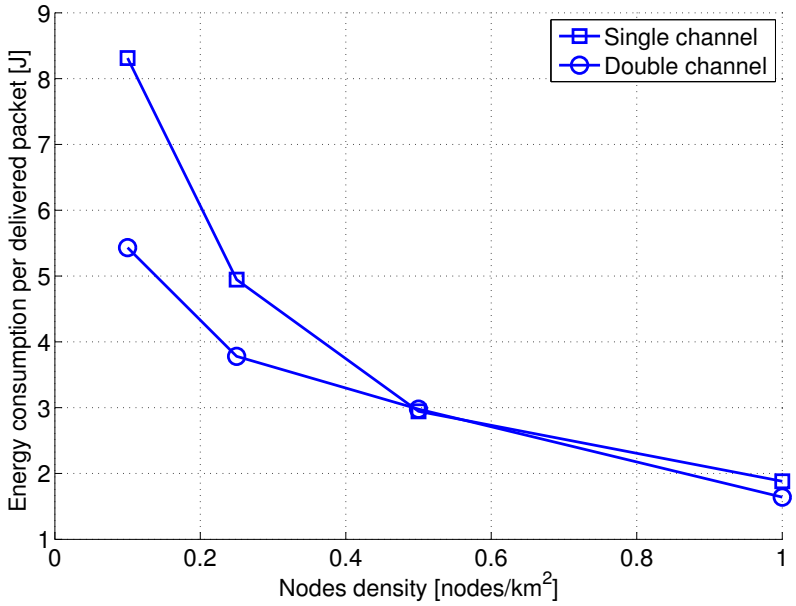


Figure 6.16: Transmission energy consumption per delivered packet with 100 nodes

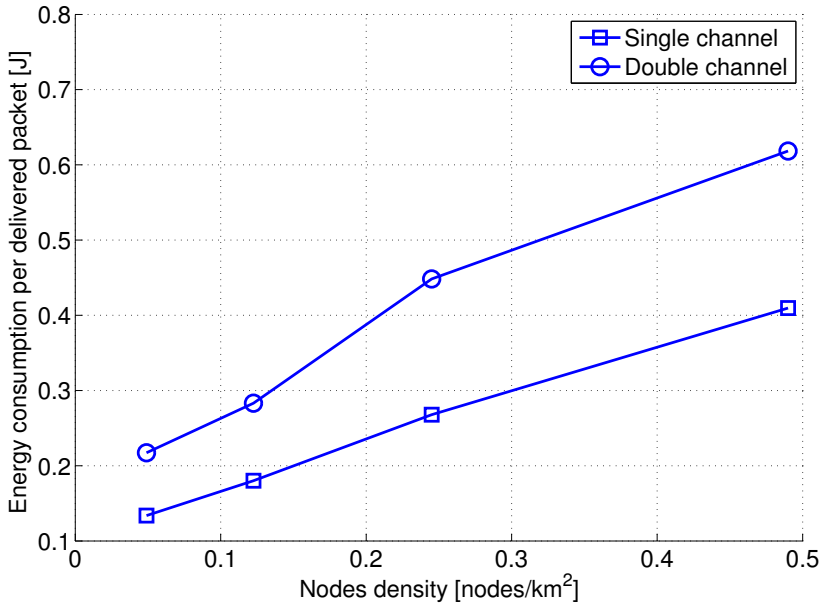


Figure 6.17: Reception energy consumption per delivered packet with 49 nodes

policy.

6.5.2 Discussion

The usage of a dynamic allocation of the bandwidth, assuming the capability of the nodes to use different power levels and the FBR routing protocol which provides control packets, allows to achieve more efficient transmission energy consumption. The actual energy efficiency of the scheme depends on the relationship between the transmission and receiving power. In fact, as a drawback, this scheme increases the receiving power due to the longer duration of some transmissions and the increased coverage range due to the better quality of the chosen channel. Moreover this scheme introduces some inefficiencies at the MAC and routing protocols.

Our scheme is principally based on the frequency selectivity of the channel, exploiting the diversity of the attenuation behavior between adjacent bandwidth. Therefore the protocol can achieve better performance at high central frequency, where the channel behavior strongly changes with the frequency. On the other hand, due to the larger transmission duration, it is necessary that power levels are widely spaced. This can be provided, by using our power level selction scheme, having a sparse network. The last point is the relationship between transmission and reception energy consumption. If the transmission power dominates the amount of energy consumption of the network, this proposed scheme can be effective for saving energy, otherwise the increased reception energy consumption cannot be balanced by

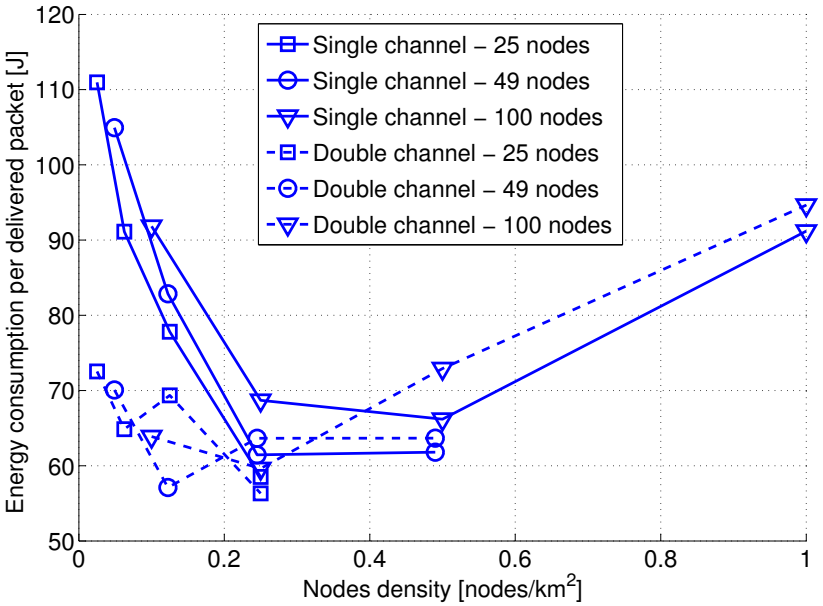


Figure 6.18: Total energy consumption per delivered packet

the more efficient transmission scheme.

An other good feature of this scheme is the simplicity. It does not require any central unit to manage the bandwidth allocation, but everything is made in a distributed manner. In this way, it is very scalable and easier to implement in a real network.

Chapter 7

Conclusion

This thesis addressed four topics: radio channel modeling, localization and traffic management in WSNs and energy efficiency in UW networks.

The approach we used to address these problems has three main aims. First, we tried to observe the problem from many points of view. For example, we proposed solutions for localization that take into account channel modeling, hardware problems, ranging techniques, beacons placement and cooperative schemes, static and mobile networks. Furthermore the sensor constraints are considered in the design of proposed algorithms. Therefore simplicity is one of the main focus of our proposals. Nontheless fair comparisons are always performed with complex techniques like Maximum Likelihood as a benchmark. Finally in the most cases we performed real experiments. This allows us to have very useful hints to tackle new research issues and to confirm our ideas and theoretical and simulative results.

In order to underdstand the channel behavior, we collected from our testbeds a large number of experimental measurements. Specifically we focused our attention on multi–channel measurements, due to the capability of the sensor nodes to use different channels. This approach led to interesting results. First, each channel has very similar statistical behavior compared to the other in terms of path loss exponent and shadowing variance, but if the received energy is averaged over several RF channels, the randomness of the measurements decreases. The performance depends on the number of considered RF channels, but it also strongly depends on the frequency separation between the RF channels. If two adjacent channels are considered, the variance of the shadowing power does not decrease as much as when two RF channels are far apart.

Another important insight we gained from experiments is the important effect of the anisotropy of the antenna of the sensor devices. Usually the randomness in RSSI measurements is ascribed to channel fluctuations or obstacles that reflect the radio signal. In our experiment we found that the randomness of the RSSI is mostly due to antenna pattern especially for distances shorter than 2 meters, where the LOS path is dominant. Therefore,

having the knowledge of the relative angle among nodes, it is possible to know the attenuation due to the antenna and the distance estimation from RSSI can be definitely more accurate.

These two methods to reduce the variance of the shadowing random effect can be very useful for the RSSI-based localization, leading to better position estimates.

We analyzed different beacon–based localization algorithms in static networks. We increased the number of beacons in the network to understand the actual limits of these techniques. Results showed that with few beacons, e.g. less than 5-6, a very simple algorithm like Min-Max outperforms the Maximum Likelihood approach that, due to the poor distance estimations from the channel measurements, needs more beacons to give good results.

Then we investigated the effect of beacon positions on the localization performance using both RSSI and ToA, in order to identify the deployment that yields the minimum average mean square localization error, given the area, the statistic of the spatial distribution of the nodes and the number of beacons. We observed that the position of the beacons can actually strongly affect the average localization accuracy. Finding the optimal beacon positioning is a very complex task, therefore we proposed a heuristic algorithm that leads to closed to optimal results. We compared the localization accuracy using both RSSI and ToA ranging estimation, with the beacons placed using our algorithm and a geometric deployment inside the area. Performance with the optimal placement is better using RSSI technique in a scenario with quite few beacons and non uniform distribution of nodes. The precision of ToA ranging does not depend on the distance between the two nodes, therefore a geometric placement of beacons overlaps the optimal one.

After the study of static networks, the next step was to introduce an AMR in the network. This mobile device is capable to perform self–localization with the onboard odometry, but it can also interact with the static sensor network. In addition, the AMR does not have constraints in terms of processing, energy and storage. Therefore it is feasible to design more complex algorithms, e.g., Extended Kalman Filters, to get better results. In our study we focused on the Simultaneous Localization and Mapping, an algorithm that consists in creating the map of an area without any a priori knowledge of the environment, while localizing the nodes in the WSN, by combining the information provided by the WSN and by the AMR odometry. Using our algorithm in real experiment, the AMR achieves quite accurate positioning estimates. In fact, our results are similar to previous works, although obtained in a more hostile scenario like indoor environment. Sensors have benefit from this scenario, largely improving their localization estimation compared to a static scenario. This improvement is due to the large number of *virtual* beacons sent by the AMR and the accuracy of the odometry of the robot compared to RSSI ranging.

Finally we considered an opportunistic network. The scenario consists of mobile nodes that can interact with other nodes that occasionally happen to be in proximity. Nodes have different self–localization capability and they can exchange their positioning informa-

tion during an opportunistic meeting. First, we proposed a complete framework for the meeting model, to have analytical results on the statistic of the meeting between two nodes. Using this model, we ran two different algorithms to manage and exploit this additional information, one based on Maximum Likelihood and a very simple heuristic, called GOAL. We compared the two techniques, finding that if the scenario is not very harsh, GOAL algorithm performs very well, slightly better than ML, while the statistical approach is very robust and improves localization accuracy even in difficult scenarios. The entire system was compared with experimental data provided by our testbed. The collected data confirmed our theoretical model giving very similar results. Therefore a node, which has a very good localization accuracy, can be used to help low–cost devices with poor localization to improve their position estimate by means of a very simple technique.

In the same scenario, we proposed another solution using the Linear Matrix Inequality and a barycentric algorithm. A fixed node has no information on his position and infers it by means of the interaction with other nodes that enter his coverage range. We implemented the analytical framework for the meeting model. We evaluated this algorithm under different scenarios, changing speed of nodes and localization accuracy of the mobile nodes. Performance are slightly affected by these parameters. In fact, the accuracy of the localization estimation is mainly affected by the number of meetings and the coverage range.

The second problem addressed in this thesis regards the traffic management in WSN. Usually all the nodes in a WSN send packets to the sink. Because all the traffic flows converge to a single node, in presence of a certain traffic load the network can suffer congestion problem, losing reliability, due to packet losses and high delay, and wasting energy. Our proposed protocol, called EPC^2, mitigates this problem improving both throughput and energy efficiency. The drawback is a higher delay that, however can be tolerated in many WSN applications. This solution is suitable for real implementation because the only additional complexity consists in syncronization, but limited only a few nodes, namely the sink and his neighbors.

Another scenario we look into is that of underwater wireless sensor networks (UWSN). Specifically we addressed the energy efficiency problem, studying two different schemes. Nodes have the capability to switch off the acoustic modem following a deterministic duty–cycle and they can perform power control to save transmission energy. Results showed that this scheme brings some benefit only in scenarios when transmissions are not the main source of energy consumption. The second proposal was an optimal management of the bandwidth. The main idea is based on the capability of the nodes to use power control and the different channel response changing the frequency. This scheme can achieve better results as the node density decreases and the central frequency increases. Also in this case the benefit of this approach depends on the balance among the node power consumption in the different states. Another characteristic is the simplicity, taking into account the hardware constraints

of underwater sensor nodes. Both solutions do not require additional hardware or centralized management.

Considering our research on localization problem it is possible to define some important results. First, the problem can be addressed from many points of view. Therefore, it is possible to combine different solutions to sum up the single contributions to improve accuracy. Moreover localization performance can benefit of mobility and heterogeneity. The proposed opportunistic localization is a novel paradigm that can lead to a number of applications and new protocols. Our simple approach yields very interesting results and can be used as a starting point for developing new schemes.

Comparison between radio and underwater network is also very interesting. The differences in propagation characteristics, network requirements, hardware constraints and bandwidth availability make well known protocols for radio network not always effective in underwater environment. To this end also simple techniques adapted to the new scenario, such duty–cycle in our case, can provide interesting results and hints for open issues. Therefore researchers have lot of possibilities to adapt radio algorithms or propose new ones in the future to improve underwater communications.

Bibliography

[1] M. Zorzi, "A new contention-based MAC protocol for geographic forwarding in ad-hoc and sensor networks," *IEEE International Conference on Communications*, vol. 6, pp. 3481–3485, Jun 2004.

[2] M. Zorzi and R. Rao, "Geographic random forwarding (GeRaF) for ad hoc and sensor networks: multihop performance," *IEEE Trans. on Mobile Computing*, vol. 2, pp. 349–364, Oct-Dec 2003.

[3] G. Zanca, F. Zorzi, A. Zanella, and M. Zorzi, "Experimental comparison of RSSI-based localization algorithms for indoor wireless sensor networks," in *In Proceedings of ACM Workshop of Real-world Wireless Sensor Networks - Glasgow*, April 2008.

[4] E. Menegatti, A. Zanella, S. Zilli, F. Zorzi, and E. Pagello, "Range-only slam with a mobile robot and a wireless sensor network," in *In Proceedings of IEEE ICRA Conference - Kobe*, 2009.

[5] F. Zorzi and A. Zanella, "Opportunistic localization: modeling and analysis," in *Proceedings of VTC Spring - Barcelona*, February 2009.

[6] F. Zorzi, A. Bardella, and A. Zanella, "Theoretical and experimental analysis of opportunistic localization algorithms in wireless networks - UNDER REVIEW AT IEEE SECON 2010."

[7] F. Zorzi, G. Kang, T. Pérennou, and A. Zanella, "Opportunistic localization scheme based on linear matrix inequality," in *WISP 2009: Proceedings of IEEE International Symposium on Intelligent Signal Processing - Budapest, Hungary*, August 2009.

[8] F. Zorzi, A. Bardella, G. Kang, T. Pérennou, and A. Zanella, "Analysis of opportunistic localization algorithms based on the linear matrix inequality method," in *Proceedings of The Second International Workshop on Mobile Opportunistic Networking ACM/SIGMOBILE MobiOpp*, February 2010.

[9] F. Zorzi, G. Kang, T. Pérennou, A. Zanella, and A. Diaz, "Group behavior impact on an opportunistic localization scheme," in *under review at Future Network and Mobile Summit*, 2010.

[10] F. Zorzi, A. Bardella, G. Kang, T. Pérennou, and A. Zanella, "Experimental localization application in opportunistic scenario," in *Demo paper in The Second International Workshop on Mobile Opportunistic Networking ACM/SIGMOBILE MobiOpp*, February 2010.

[11] P. Casari, F. Zorzi, and M. Zorzi, "Efficient packet converge-casting: Relieving the sink congestion in wireless sensor networks," in *Personal, Indoor and Mobile Radio Communications, 2007. PIMRC 2007. IEEE 18th International Symposium on*, Sept. 2007, pp. 1–5.

[12] F. Zorzi, M. Stojanovic, and M. Zorzi, "On the effects of node density and duty–cycle on energy–efficiency in underwater networks," in *under review at IEEE OCEANS*, 2010.

[13] ——, "An adaptive channel management scheme for energy–efficiency in underwater networks," in *under review at IEEE OCEANS*, 2010.

[14] T. S. Rappaport, *Wireless Communications: Principles and Practice.* Piscataway, NJ, USA: IEEE Press, 1996.

[15] A. Goldsmith, *Wireless Communications.* New York, NY, USA: Cambridge University Press, 2005.

[16] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, "Locating the nodes: cooperative localization in wireless sensor networks," *Signal Processing Magazine, IEEE*, vol. 22, no. 4, pp. 54–69, July 2005.

[17] Infineon, "EyesIFX sensor network development kit- documentation." [Online]. Available: www.infineon.com

[18] R. Crepaldi, A. Harris, A. Scarpa, A. Zanella, and M. Zorzi, "Signetlab: deployable sensor network testbed and management tool," in *SenSys '06: Proceedings of the 4th international conference on Embedded networked sensor systems.* New York, NY, USA: ACM Press, 2006, pp. 375–376.

[19] ——, "Testbed implementation and refinement of a range-based localization algorithm for wireless sensor networks," in *3rd IEE Mobility Conference 2006.*, Oct., 25-27. 2006.

[20] "Tmote sky datasheet," MoteIv Corporation. [Online]. Available: www.moteiv.com

[21] G. K. Kraetzschmar, H. Utz, S. Sablatnög, S. Enderle, and G. Palm, "Miro - middleware for cooperative robotics," in *RoboCup 2001: Robot Soccer World Cup V*. London, UK: Springer-Verlag, 2002, pp. 411–416.

[22] H. Utz, S. Sablatnog, S. Enderle, and G. Kraetzschmar, "Miro - middleware for mobile robot applications," *Robotics and Automation, IEEE Transactions on*, vol. 18, no. 4, pp. 493–497, Aug 2002.

[23] A. Savvides, H. Park, and M. Srivastava, "The bits and flops of the N-hop multilateration primitive for node localization problems," in *First ACM International Workshop on Wireless Sensor Networks and Application*, September 2002. [Online]. Available: citeseer.ist.psu.edu/savvides02bits.html

[24] M. Banatre, A. Ollero, A. Wolisz, and P. J. Marron, *Cooperating Objects and Wireless Sensor Networks*. Hermes Science, 2007.

[25] C. Wang and L. Xiao, "Sensor localization under limited measurement capabilities," *IEEE Network*, vol. 21, no. 3, pp. 16–23, May-June 2007.

[26] A. Savvides, H. Park, and M. B. Srivastava, "The N-hop multilateration primitive for node localization problems," *Mobile Network Applicatons*, vol. 8, no. 4, pp. 443–451, 2003.

[27] N. Patwari, R. O'Dea, and Y. Wang, "Relative location in wireless networks," in *Vehicular Technology Conference, 2001. VTC 2001 Spring. IEEE VTS 53rd*, vol. 2, 2001, pp. 1149–1153 vol.2.

[28] R. P. M. Eiman Elnahrawy, Xiaoyan Li, "The limits of localization using signal strength: a comparative study," in *Proceeding of Sensor and Ad Hoc Communications and Networks, IEEE SECON*, October 2004, pp. 406–414.

[29] T. He, C. Huang, B. M. Blum, J. A. Stankovic, and T. Abdelzaher, "Range-free localization schemes for large scale sensor networks," in *MobiCom '03: Proceedings of the 9th annual international conference on Mobile computing and networking*. New York, NY, USA: ACM, 2003, pp. 81–95.

[30] C. Savarese, J. M. Rabaey, and K. Langendoen, "Robust positioning algorithms for distributed ad-hoc wireless sensor networks," in *ATEC '02: Proceedings of the General Track of the annual conference on USENIX Annual Technical Conference*. Berkeley, CA, USA: USENIX Association, 2002, pp. 317–327.

[31] P. Bahl and V. Padmanabhan, "Radar: an in-building rf-based user location and tracking system," *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 2, pp. 775–784 vol.2, 2000.

[32] K. Langendoen and N. Reijers, "Distributed localization in wireless sensor networks: a quantitative comparison," *Comput. Networks*, vol. 43, no. 4, pp. 499–518, 2003.

[33] T. R. XuanLong Nguyen, "Localization algorithms for sensor networks using rf signal strength CS 252 class project," May 2003.

[34] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.

[35] T. H. Chong Liu, Kui Wu, "Sensor localization with ring overlapping based on comparison of received signal strength indicator," *IEEE Trans. on Mobile Computing*, 2004.

[36] N. Bulusu, D. Estrin, and J. Heidemann, "Adaptive beacon placement," *icdcs*, vol. 00, p. 0489, 2001.

[37] D. Kurth, G. Kantor, and S. Singh, "Experimental results in range-only localization with radio," *Intelligent Robots and Systems, 2003. (IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on*, vol. 1, pp. 974–979 vol.1, Oct. 2003.

[38] J. Djugash, S. Singh, G. Kantor, and W. Zhang, "Range-only slam for robots operating cooperatively with sensor networks," *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*, pp. 2078–2084, 15-19, 2006.

[39] P. Corke, S. Hrabar, R. Peterson, D. Rus, S. Saripalli, and G. Sukhatme, "Autonomous deployment and repair of a sensor network using an unmanned aerial vehicle," *Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on*, vol. 4, pp. 3602–3608 Vol.4, 26-May 1, 2004.

[40] E. Fasolo, C. Prehofer, M. Rossi, Q. Wei, J. Widmer, A. Zanella, and M. Zorzi, "Energy-efficiency: challenges and new approaches for efficient data gathering and dissemination in pervasive wireless networks," in *Proc. of INTERSENSE06*, 2006.

[41] N. B. Priyantha, A. Chakraborty, and H. Balakrishnan, "The cricket location-support system," in *MobiCom '00: Proceedings of the 6th annual international conference on Mobile computing and networking*. New York, NY, USA: ACM, 2000, pp. 32–43.

[42] G. M. Rebeiz, *RF MEMS: Theory, Design, and Technology*. New York, NY, USA: John Wiley & Sons, Inc., 2003.

[43] F. V. Diggelen and C. Abraham, "Indoor gps technology," *CTIA Wireless Agenda - Dallas, May 2001*, 2001.

[44] L. Doherty, L. E. Ghaoui, and K. S. J. Pister, "Convex position estimation in wireless sensor networks," in *Proc. IEEE Infocom*, Anchorage, AK, Apr. 2001, pp. 1655–1663, http://www-bsac.EECS.Berkeley.EDU/~ldoherty/infocom.pdf.

[45] J. Leonard and H. Durrant-Whyte, "Mobile robot localization by tracking geometric beacons," *Robotics and Automation, IEEE Transactions on*, vol. 7, no. 3, pp. 376–382, Jun 1991.

[46] F. Dellaert, D. Fox, W. Burgard, and S. Thrun, "Monte carlo localization for mobile robots," *Robotics and Automation, 1999. Proceedings. 1999 IEEE International Conference on*, vol. 2, pp. 1322–1328 vol.2, 1999.

[47] D. Fox, W. Burgard, F. Dellaert, and S. Thrun, "Monte carlo localization: efficient position estimation for mobile robots," in *AAAI '99/IAAI '99: Proceedings of the sixteenth national conference on Artificial intelligence and the eleventh Innovative applications of artificial intelligence conference*. Menlo Park, CA, USA: American Association for Artificial Intelligence, 1999, pp. 343–349.

[48] D. Fox, W. Burgard, H. Kruppa, and S. Thrun, "A probabilistic approach to collaborative multi-robot localization," *Autonomous Robots*, vol. 8, pp. 325–344(20), June 2000.

[49] S. Roumeliotis and G. Bekey, "Collective localization: a distributed kalman filter approach to localization of groups of mobile robots," *Robotics and Automation, 2000. Proceedings. ICRA '00. IEEE International Conference on*, vol. 3, pp. 2958–2965 vol.3, 2000.

[50] A. Howard, M. Matark, and G. Sukhatme, "Localization for mobile robot teams using maximum likelihood estimation," *Intelligent Robots and System, 2002. IEEE/RSJ International Conference on*, vol. 1, pp. 434–439 vol.1, 2002.

[51] H. Lee and H. Aghajan, "Collaborative node localization in surveillance networks using opportunistic target observations," in *VSSN '06: Proceedings of the 4th ACM international workshop on Video surveillance and sensor networks*. New York, NY, USA: ACM, 2006, pp. 9–18.

[52] J. M. Roberts, E. S. Duff, and P. I. Corke, "Reactive navigation and opportunistic localization for autonomous underground mining vehicles," *Information Sciences*, vol. 145, pp. 127–146(20), August 2002.

[53] M. Musolesi, S. Hailes, and C. Mascolo, "Adaptive routing for intermittently connected mobile ad hoc networks," *A World of Wireless, Mobile and Multimedia Networks, International Symposium on*, vol. 1, pp. 183–189, 2005.

[54] A. Vahdat and D. Becker, "Epidemic routing for partially-connected ad hoc networks," Duke Tech Report CS-2000-06, Tech. Rep., 2000.

[55] C. Boldrini, M. Conti, I. Iacopini, and A. Passarella, "Hibop: a history based routing protocol for opportunistic networks," in *Proc. IEEE WoWMoM 2007*, 2007.

[56] M. Conti, A. Passarella, and L. Pelusi, "Mobile-relay forwarding in opportunistic networks," in *Chapter in Adaptive Techniques in Wireless Networks (M. Ibnkahla, Editor.* CRC Press, 2008.

[57] S. Jain, K. Fall, and R. Patra, "Routing in a delay tolerant network," in *SIGCOMM '04: Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications.* New York, NY, USA: ACM, 2004, pp. 145–158.

[58] IEEE Standards Department, *ANSI / IEEE Standard 802.11.* IEEE Press, 1999.

[59] C. Y. Wan, S. B. Eisenman, and A. T. Campbell, "Coda: congestion detection and avoidance in sensor networks," in *Proc. of ACM SenSys*, Los Angeles, CA, nov 2003, pp. 266–279.

[60] O. B. Akan and I. F. Akyildiz, "Event–to–sink reliable transport in wireless sensor networks," *IEEE_J_NET*, vol. 13, pp. 1003–1016, oct 2005.

[61] C. T. Ee and R. Bajcsy, "Congestion control and fairness for many-to-one routing in sensor networks," in *Proc. of ACM SenSys*, Baltimore, MD, nov 2004, pp. 148–161.

[62] M. Vuran, V. Gungor, and Ö. B. Akan, "On the interdependence of congestion and contention in wireless sensor networks," in *Proc. of SenMetrics*, San Diego, CA, jul 2005, pp. 136–147.

[63] F. 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_J_COM*, vol. 23, no. 12, pp. 1417–1433, dec 1975.

[64] J. G. Proakis, *Digital Communications*, 2nd ed. New York: McGraw-Hill, 1999.

[65] B. Peleato and M. Stojanovic, "A MAC protocol for ad-hoc underwater acoustic sensor networks," in *WUWNet '06: Proceedings of the 1st ACM international workshop on Underwater networks.* New York, NY, USA: ACM, 2006, pp. 113–115.

[66] M. K. Park and V. Rodoplu, "UWAN-MAC: An energy-efficient MAC protocol for underwater acoustic wireless sensor networks," *Oceanic Engineering, IEEE Journal of*, vol. 32, no. 3, pp. 710–720, July 2007.

[67] A. Harris and M. Zorzi, "On the design of energy-efficient routing protocols in underwater networks," in *Sensor, Mesh and Ad Hoc Communications and Networks, 2007. SECON '07. 4th Annual IEEE Communications Society Conference on*, June 2007, pp. 80–90.

[68] P. Casari, M. Stojanovic, and M. Zorzi, "Exploiting the bandwidth-distance relationship in underwater acoustic networks," in *OCEANS 2007*, 29 2007-Oct. 4 2007, pp. 1–6.

[69] M. Stojanovic, "On the relationship between capacity and distance in an underwater acoustic communication channel," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 11, no. 4, pp. 34–43, 2007.

[70] J. M. Jornet, M. Stojanovic, and M. Zorzi, "Focused beam routing protocol for underwater acoustic networks," in *WuWNeT '08: Proceedings of the third ACM international workshop on Underwater Networks*. New York, NY, USA: ACM, 2008, pp. 75–82.

[71] A. F. Harris, III, M. Stojanovic, and M. Zorzi, "When underwater acoustic nodes should sleep with one eye open: idle-time power management in underwater sensor networks," in *WUWNet '06: Proceedings of the 1st ACM international workshop on Underwater networks*. New York, NY, USA: ACM, 2006, pp. 105–108.

[72] L. M. Brekhovskikh and Y. P. Lysanov, "Fundamentals of ocean acoustics (3rd edition)," *The Journal of the Acoustical Society of America*, vol. 116, no. 4, pp. 1863–1863, 2004. [Online]. Available: http://link.aip.org/link/?JAS/116/1863/1

[73] R. Coates, *Underwater Acoustic Systems*. New York: Wiley, 1989.

[74] J. Montana, "AUVNetsim: A simulator for underwater acoustic networks," MIT Sea Grant Technical Report, 2008. [Online]. Available: http://users.ece.gatech.edu/jmjm3/publications/auvnetsim.pdf

List of Publications

Accepted and published papers

A1. P. Casari, F. Zorzi, and M. Zorzi, "Efficient packet converge-casting: Relieving the sink congestion in wireless sensor networks," in *IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sept. 2007.

A2. G. Zanca, F. Zorzi, A. Zanella, and M. Zorzi, "Experimental comparison of RSSI-based localization algorithms for indoor wireless sensor networks," in *In Proceedings of ACM Workshop of Real-world Wireless Sensor Networks - Glasgow*, April 2008.

A3. F. Librino, S. Del Favero, F. Zorzi, A. F. Harris III and M. Zorzi, "A Distributed Solution to Estimation Problems in Wireless Sensor Networks leveraging Broadcast Communications," in *Proceedings of IEEE WONS Conference - Snowbird UTAH*, February 2009.

A4. F. Zorzi and A. Zanella, "Opportunistic localization: modeling and analysis," in *Proceedings of VTC Spring - Barcelona*, February 2009.

A5. E. Menegatti, A. Zanella, S. Zilli, F. Zorzi, and E. Pagello, "Range-only slam with a mobile robot and a wireless sensor network," in *In Proceedings of IEEE ICRA Conference - Kobe*, May 2009.

A6. F. Zorzi, G. Kang, T. Pérennou, and A. Zanella, "Opportunistic localization scheme based on linear matrix inequality," in *WISP 2009: Proceedings of IEEE International Symposium on Intelligent Signal Processing - Budapest, Hungary*, August 2009.

A7. F. Zorzi, A. Bardella, G. Kang, T. Pérennou, and A. Zanella, "Analysis of opportunistic localization algorithms based on the linear matrix inequality method," in *Proceedings of The Second International Workshop on Mobile Opportunistic Networking ACM/SIGMOBILE MobiOpp*, February 2010.

A8. F. Zorzi, A. Bardella, G. Kang, T. Pérennou, and A. Zanella, "Experimental localization application in opportunistic scenario," *Demo paper in The Second International Workshop on Mobile Opportunistic Networking ACM/SIGMOBILE MobiOpp*, February 2010.

Submitted papers

S1. F. Zorzi, A. Bardella, and A. Zanella, "Theoretical and experimental analysis of opportunistic localization algorithms in wireless networks," *under review at IEEE SECON*, 2010.

S2. F. Zorzi, G. Kang, T. Pérennou, A. Zanella, and A. Diaz, "Group behavior impact on an opportunistic localization scheme," *under review at Future Network and Mobile Summit*, 2010.

S3. F. Zorzi, M. Stojanovic, and M. Zorzi, "On the effects of node density and duty–cycle on energy–efficiency in underwater networks," *under review at IEEE OCEANS*, 2010.

S4. F. Zorzi, M. Stojanovic, and M. Zorzi, "An adaptive channel management scheme for energy–efficiency in underwater networks," *under review at IEEE OCEANS*, 2010.