

RF-BASED LOCALIZATION IN GPS-DENIED APPLICATIONS

by

Gianni Giorgetti

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

ARIZONA STATE UNIVERSITY

August 2009

RF-BASED LOCALIZATION IN GPS-DENIED APPLICATIONS

by

Gianni Giorgetti

has been approved

August 2009

Graduate Supervisory Committee:

Sandeep Gupta, Chair
Karamvir Chatha
Tolga Duman
Stephen Goodnick
Kari Torkkola

ACCEPTED BY THE GRADUATE COLLEGE

ABSTRACT

Recent years have witnessed the emergence of novel application paradigms such as the Wireless Sensor Network and Context Aware computing. Among the challenges posed by these applications, localization – i.e. the process of locating people and/or devices – has emerged as a key problem that has found only partial answers. Although GPS receivers are common on many consumer electronic devices, alternative solutions are needed when locating devices that strive to be small and inexpensive, as in sensor networks, or when supporting indoor positioning. This dissertation focuses on radio-based positioning schemes suitable for applications where GPS is not a viable solution.

The first part of this work addresses schemes that use proximity constraints inferred from radio connectivity. A novel solution based on the Self-Organizing Map (SOM) formalism is proposed. Using extensive simulations, the SOM approach is shown to achieve a low localization error using limited computational resources. Comparison with other schemes demonstrate favorable results, especially in sparse deployments and when few (or none) of the nodes are located at known positions.

The second part focuses on theoretical analysis of the results. Two broad families of positioning schemes are analyzed: 1) Range-free schemes that use radio proximity information, as in the SOM approach; and 2) Range-based schemes that measure the attenuation of the Radio-Frequency (RF) signal to estimate inter-node distances. First, analysis of the Fisher Information and the Cramér–Rao bound are used to investigate the theoretical limits that bound the localization error in the two cases. Then, general design criteria are proposed to reduce the error of range-free schemes and determine in which operative conditions they can outperform range-based solutions.

In the final part of this work, the theoretical results are used to design an improved variant of the SOM algorithm that combines the best traits of proximity and RSS ranging localization. Validation using measurements from real deployments shows significant improvements over the original SOM version and other localization schemes. Practical implementation of RF-based positioning systems is further investigated by using directional antennas for Angle of Arrival (AOA) estimation. A novel angle-based system that uses a single anchor is described and validated using experimental results. Additionally, a SOM variant capable of exploiting AOA information in collaborative localization is investigated using simulations.

ai miei genitori

(to my parents)

ACKNOWLEDGMENTS

This dissertation concludes the work performed during a joint doctoral program between the Università degli Studi di Firenze and Arizona State University, a course of studies that has allowed me to meet many great people both in Italy and in the U.S. I owe sincere gratitude to Gianfranco Manes and Stephen Goodnick for making this program possible.

I am indebted to my advisor Sandeep Gupta for his support and guidance. He has provided me with valuable advice and encouragement during my time of study at Arizona State. I also want to thank my committee members and all the great teachers from who I have learned so much during the past few years.

I am grateful to Alessandro Cidronali, Stefano Maddio, and Stefano Maurri from the Microelectronic Lab at the University of Florence. They designed and built the antenna prototypes used in the last part of this dissertation. Without their enthusiastic work, part of my research would have not been possible. I also owe a big thank-you to my student colleagues at IMPACT lab for making my stay in AZ particularly enjoyable: Guofend Deng, Su-Jin Kim, Tridib Mukherjee, Qinghui Tang, Krishna Venkatasubramanian, Edward Raleigh, Ken Bannister, and Georgios Varsamopoulos. I also thank NSF, Intel Corp., and the *Ministero dell'Istruzione, dell'Università e della Ricerca* (MIUR) for funding part of my research work.

Finally, my deepest gratitude goes to my family: my father, my mother, and my two sisters Laura and Monica. They all encouraged me during these long years of studies. Without their warm support this work would have not been possible.

TABLE OF CONTENTS

Chapter	Page
LIST OF TABLES	1
LIST OF FIGURES	1
1 Introduction	1
1.1 Wireless Sensor Networks and Indoor Applications	1
1.2 Collaborative Localization	2
1.3 Outline and Contributions	4
1.3.1 Range-Free Localization	4
1.3.2 The Limits Of Radio-Based Localization	5
1.3.3 Implementation of RF-Based Positioning Systems	6
2 Background	8
2.1 System Model and Problem Definition	8
2.1.1 Assumptions	10
2.2 The complexity of the Node Positioning Problem	11
2.2.1 Embedding With Known Edge Lengths	11
2.2.2 Embedding Using Connectivity Information	14
2.3 Localization Schemes in Wireless Sensor Networks	15
2.3.1 Range-Based Algorithms	16
2.3.2 Range-Free Algorithms	18
2.3.3 Scene Analysis Algorithms	24
3 Range-Free Localization Using Self-Organizing Maps	25
3.1 The SOM Learning Algorithm	25
3.1.1 Learning Parameters	27
3.1.2 Properties	28
3.2 Localization using SOMs	28

Chapter	Page
3.2.1 System Model	30
3.2.2 Modified SOM Model	30
3.2.3 Localization Algorithm	31
3.3 Simulation Model	32
3.3.1 Placement Model	33
3.3.2 Connectivity Model	33
3.3.3 Error Metric	34
3.4 Anchor-Free Localization: Virtual Coordinates and Geo-Routing	34
3.5 Anchor-Based Localization: Absolute Coordinates	36
3.5.1 Exploiting Anchor Information: The SOM-A Scheme	36
3.5.2 Comparison Between SOM-A and SOM-V	37
3.5.3 Weight Initialization and Convergence	38
3.5.4 Comparison With MDS and DV-HOP	40
3.6 Computational Complexity Analysis	42
3.7 Implementation on a Resource-Constrained Sensor Node	43
3.8 Online vs Batch Training	44
3.9 Comparison with Distributed Localization Schemes	45
3.10 Related Work	47
3.10.1 Localization Using SOM	47
3.10.2 Theoretical results	48
3.10.3 Discussion	49
4 Understanding the Limits of RF-Based Localization	50
4.1 Localization Based on Radio Connectivity	50
4.1.1 Motivating Example	51
4.1.2 Range-Free Localization as a Parameter Estimation Problem	53
4.1.3 Preliminaries on the Parameter Estimation Approach	54
4.1.4 Single Node Localization	55
4.1.5 The log-normal shadowing model	55
4.1.6 Threshold-based connectivity	56

Chapter	Page
4.1.7 Fisher Information and Cramér-Rao Bound Analysis	57
4.1.8 Optimal Threshold for the 1D Case	59
4.2 Optimal Threshold Selection in Collaborative Localization: The Optimal Connectivity (OC) Value	61
4.2.1 CRB Analysis	61
4.2.2 The Optimal Connectivity (OC) Value	64
4.2.3 Properties of the OC Value.	68
4.2.4 Approximation of the Optimal Connectivity Value	74
4.2.5 Test Cases	77
4.3 Comparison with the RSS-Ranging Approach	79
4.3.1 1D Node Range-Based Localization	79
4.3.2 To Range Or Not To Range?	81
4.3.3 Effect of Shadowing on Range-Free and Range-Based Localization	82
4.3.4 Network Localization	83
4.4 Range-Free and Range-Based Collaborative Localization: The Critical Connectivity (CC) Value	83
4.4.1 Properties of the Critical Connectivity	84
4.4.2 Critical Connectivity Approximation	86
4.4.3 Test Case	88
4.5 Related Work	89
5 Localization in Heterogeneous Scenarios Using SOM	91
5.1 Performance of the SOM based Localization under Log-Normal Shadowing and Comparison with the CRB	92
5.1.1 Localization in 1D and 3D spaces	94
5.2 Localization in Dense Networks	94
5.2.1 The SOM-R algorithm	96
5.2.2 Simulation Results	98
5.2.3 Localization Using RSS Data From Real Deployments	101
5.3 Localization in Anisotropic Deployments	103

Chapter	Page
6 Localization using Directional Antennas	106
6.1 Azimuthal Angle of Arrival Estimation	108
6.1.1 Four Beam Directional Antenna (FBDA)	108
6.1.2 Principle Of Operations	109
6.1.3 AOA Estimation	111
6.1.4 AOA Estimation Results using In-Field Measurement Data	114
6.2 Indoor localization Using a Single Anchor Node	116
6.2.1 Antenna Design	116
6.2.2 Localization Application	117
6.2.3 Principle of operations	117
6.2.4 Range-Free Localization (Proximity)	118
6.2.5 Range-Based Localization (DOA Estimation)	121
6.2.6 Fingerprinting	121
6.2.7 Discussion	122
6.3 Use of Angle Information in Collaborative Localization Schemes	123
6.3.1 The SOM Localization Scheme Using AOA Information	124
6.3.2 Collaborative Single-Anchor Localization	125
6.3.3 RSS-Based Collaborative Localization	128
7 Conclusions	130
7.1 Future Work	131
References	135

LIST OF TABLES

Table	Page
3.1 Memory requirements and execution time of the SOM-A algorithm on a TelosB node equipped with a 8 MHz microcontroller.	44
4.1 Effect of Parameters on CC and OC Values	86
5.1 Localization results in anisotropic networks.	105

LIST OF FIGURES

Figure	Page
1.1 Collaborative localization in ad-hoc wireless networks. The node positions are computed using information collected by neighboring devices.	3
2.1 A centralized localization system. The unknown node positions are computed using information about the number of nodes n in the system, the position $\{[x_a, y_a]^t\}$ of the anchors (if available), and the set of measurements \mathcal{M} collected between neighboring nodes.	9
2.2 Localization ambiguities in absence of reference points.	12
2.3 Flex Ambiguities: a) Node 5 can be moved continuously along a circular path; b) Node 2 admits two positions compatible with the measured distances from nodes 1 and 5.	12
2.4 Discontinuous flex ambiguities. If edge 2-6 is removed, the graph can be deformed to obtain a new realization where edge 2-6 can be reinserted without changing its length.	13
2.5 A Unit Disk Graph (UDG) used to model a network with connectivity constraints.	15
2.6 Localization in 2D: a) Trilateration: distance measurements from 3 non-collinear anchor points, b) Triangulation: angle measurements from 2 anchor point, c) angle and range from a single anchor point	16
2.7 Centroid scheme: the position of the node to localize is given by the center of gravity (COG) of anchors heard.	18
2.8 DV-Hop scheme: the shortest path (hop-count, hc) is used to estimate the distance from a node x to the anchor nodes in the network. The actual position is computed using multilateration.	20
2.9 APIT scheme: a) Each node uses the APIT test to determine if it is inside the triangle area defined by three anchors. b) The final position is the COG of the intersection of all triangles a node belong to.	21
2.10 Localization error of different range-free schemes reported by He et al. [64].	21
2.11 SeRLoc scheme: the anchor nodes generate <i>sectorized</i> beacons using directional antennas. The final node position is the COG of intersection area.	22
2.12 Lighthouse: a parallel beam is generated by a rotating base station. Nodes determine the distance from the base station by timing the light beam.	23

Figure	Page
2.13 SpotLight: localization is implemented by timing the arrive of a light beam. Three options are available: Point Scan, Line Scan, and Area Scan.	24
3.1 a) Self-Organizing Map with hexagonal pattern; b, c) two steps of the training algorithm. . .	26
3.2 a) Exponential scheduling for the parameter σ ; b) neighborhood function at different number iterations; c) portion of a SOM showing the map distances from the BMU.	27
3.3 10×10 SOM trained with samples from the RGB color space.	28
3.4 5×5 SOM trained with random samples from a 2D training set.	30
3.5 Correspondence between the nodes in a WSN (left) and the neurons in a SOM (right). The arrangement of neurons in the SOM describes the neighborhood relationship in the WSN. The weight vector of each neuron contains an estimate of the corresponding node coordinates.	31
3.6 a) The noisy grid model; b, c) two 100-node networks with different perturbation factors. . .	33
3.7 a) Ad-hoc network topology; b) Routing path discovered by a geo-routing scheme using the coordinates produced by SOM-V. Note: the two maps has been aligned using the three reference marks on the plot, but this step is not necessary for routing.	35
3.8 Average performance of a greedy geo-routing algorithm using the true coordinates and the virtual maps produced by SOM-V: a) Delivery Ratio; b) Path Length. The path length plots (right) are almost completely overlapping for the two cases.	36
3.9 Average localization error for a set of 50 random topologies with 64 nodes deployed in a $30 \text{ m} \times 30 \text{ m}$ square region.	38
3.10 Average localization error as a function of the number of iterations used in training the map. The results were generated using the same set of networks discussed in the previous section.	39
3.11 Localization example. If the neighborhood function's initial radius is not wide enough, SOM will occasionally produce incorrect results (case c). This problem is avoided by using the values in Algorithm 1.	40
3.12 Average error achieved by SOM-A, DV-HOP, and MDS in localizing sets of 50 networks with increasing perturbation factors. All networks have four anchors.	41
3.13 Average error of SOM, DV-HOP, and MDS in localizing sets networks with increasing number of nodes and four anchors: a) Network connectivity equal to five; b) Network connectivity equal to ten.	42

Figure	Page
3.14 Number of messages transmitted in a 400 node sensor network to enable localization with DV-HOP (distributed) and SOM (centralized). Neither flooding nor the collection tree protocol have been optimized.	45
3.15 Average number of messages per node transmitted in a sensor network to enable localization with DV-HOP (distributed) and SOM (centralized).	46
4.1 Localization errors for a 44-node network deployed in the cubicles of an office space. a) Original Network; b,c,d) Localization results using DV-HOP, MDS, and SOM. Segments of lines are used to connect the true node position to the estimated ones. The long lines in all the three plots denote large errors.	52
4.2 Localization error of SOM, MDS, and DV-HOP for the network in Figure 4.1a. The error is plotted as a function of the threshold P_{th} used to quantize the RSS data.	53
4.3 Schematic representation of the parameter estimation approach.	54
4.4 1D localization: the distance of node 1 from the origin has to be estimated using connectivity information obtained by quantization of the RSS data.	55
4.5 a) Probability density function of Z and probability p of detecting the node as “connected”; b) Probability of the event “nodes connected” as a function of d_{th} ($d = 5$ m).	57
4.6 Fisher Information as a function of the threshold distance d_{th} : a) for nodes at $d = 5$ m and various value of the ratio σ_{dB}/n_p ; b) for nodes at distance $d = \{2.5, 5.0, 7.5\}$ m and $\sigma_{dB}/n_p = 6/3$	59
4.7 Localization example with a node placed in three possible positions.	60
4.8 Distribution of the RSS values for a node at positions A, B, C, and three possible thresholds.	60
4.9 CRB and average localization error for localization in 2D and 3D spaces using three range-free schemes. NOTE: the CRB is compared to the average error only to illustrate the relation between its minimum and the error of the three schemes. Meaningful comparison between the absolute values should use the RMS localization error.	66
4.10 CRB for a 49 node network with four anchors on the corner of the deployment area. The Optimal Connectivity (■) value is marked on the plot.	67

4.11	Effect of choosing two different threshold values. The background color indicates the Fisher Information $F_{\text{con}}(d, d_{\text{th}})$ at different distances from the node in the center. Darker colors correspond to an higher information content. Increasing d_{th} increases the number of nodes whose distance is similar to d_{th} , but since $F_{\text{con}} \propto 1/d^2$, these nodes contribute individually less information.	68
4.12	a) Optimal connectivity (■) computed for different network sizes; b) Approximation of the I_r term (solid line) using an exponential function (circles).	69
4.13	Derivative of the Fisher Information in two cases: a) all the neighbors of the reference node are at distances $d_i < d_{\text{th}}$: to increase the Fisher information d_{th} needs to be decreased; b) all the neighbors are at distances $d_i > d_{\text{th}}$: to increase the Fisher information d_{th} needs to be increased.	71
4.14	a) Optimal Connectivity (■) for a 64 node topology with different ratios σ_{dB}/n_p ; b) Value of the function $F_{\text{tot}}(d_{\text{th}})$ computed for a node in the center of a 64 node topology.	72
4.15	Optimal Connectivity: a) for different scaling factors; b) for increasing number of anchors.	73
4.16	Simulation results: Optimal Connectivity for: a) 2D networks; b) 3D networks.	74
4.17	The network connectivity is increased by adding a node at distance $d_1 = d_{\text{th}}/2$. To maintain the optimality of the threshold d_{th} a number of nodes on the perimeter of the network need to be added.	75
4.18	Trend of the theoretical optimal connectivity computed using an iterative approximation.	76
4.19	Using interpolation functions to approximate the OC values.	78
4.20	a) Localization error for the 44-node 2D network in [136]; b) localization error for the 38 node 3D network and in [109].	78
4.21	Fisher Information for RSS and connectivity measurements.	81
4.22	Top: Distribution of the RSS values for nodes at positions A, B, C, and three possible thresholds. Bottom: the same distributions computed with a larger ratio σ_{dB}/n_p . Note the the probability $p_a > 0$ that results from the noise in the measurements.	82
4.23	Effect of increasing the values of the ratio σ_{dB}/n_p the I_r term (left) and on the Fisher information values computed for various threshold selection (right).	83

Figure	Page
4.24 CRBs for a 49 node network with four anchors on the corner of the deployment area. The Critical Connectivity (●) and Optimal Connectivity (■) values are marked on the plot.	84
4.25 Critical Connectivity (●): a) for different network sizes; b) for increasing values of the ratio σ_{dB}/n_p	85
4.26 Critical Connectivity: a) for different values of the scaling factor: $S = \{1, 2, 3\}$; b) for increasing number of anchor nodes.	86
4.27 Simulation results for critical connectivity values and their approximation (2D networks).	87
4.28 Approximation of the critical connectivity values and for different intervals of the values σ_{dB}/n_p	87
4.29 Simulation results for critical connectivity values and their approximation (3D networks).	88
4.30 100 node network test case.	89
5.1 a) 2D sample topology (red squares are anchor nodes); b,c,d) average RMS error achieved by SOM-A, MDS and DV-HOP for different values of the propagation model's parameters n_p and σ_{dB}	93
5.2 a,c) 1D and 3D sample topologies (red squares are anchor nodes); b,d) average RMS error achieved by SOM-A, MDS, DV-HOP for different connectivity levels and comparison with the CRB value.	95
5.3 Average RMS error achieved by SOM-A, SOM-R, MDS and DVH.	98
5.4 Simulation results for 200 and 400 node deployment for different connectivity values.	100
5.5 Average localization error computed using SOM for an increasing number of iterations.	101
5.6 The 44-node network with RSS measurements described by Patwari et al. [136] and published localization result for the same network.	102
5.7 Node deployments and localization results for test cases 2. Red squares represent the anchor nodes used.	103
5.8 Node deployment and localization results for nodes placed in a 3D network.	103
5.9 Sample anisotropic topologies. Red squares are the anchor nodes.	104
5.10 Sample results for anisotropic layouts: the SOM-A algorithm reduces the average localization error of 75% with respect to MDS.	105

Figure	Page
6.1 One of the switched-beam antennas developed in collaboration with Microelectronics Lab at the Università Degli Studi di Firenze, Florence, Italy.	107
6.2 Azimuth AOA estimation.	108
6.3 Radiation patterns of the four antenna faces. The patch used for transmission/reception is selected using two digital lines.	109
6.4 Distances and angles of a target node relative to patch 1 and 2.	110
6.5 Distances and angles of a target node relative to patches 1-3, and patches 1-4.	111
6.6 MATLAB Code for the implementation of the music algorithm	113
6.7 Music spectrum computed from measured data for the target at 0° , 90° , 180° and 270°	113
6.8 Result of AOA estimation experiments: the table b and c report the estimation error in two cases (A and B) and using two transmission levels (-25 dBm and -15 dBm).	115
6.9 a) Bottom-up view of the switched-beam directional antenna used to implement the <i>single-anchor</i> localization system. b) Simulated radiation patterns when faces 1 and 2 are active.	116
6.10 Deployment area and 6×4 measurement grid. Both the BS and the target use a TI CC2420 transceivers set to transmit at -15 dBm.	118
6.11 Top: antenna gains of faces 1 to 6 at different locations of the deployment area. Bottom: measured RSS values from the six antenna faces on the 6×4 grid. To facilitate comparison with the gain values, the RSS values are centered on their mean and normalized in the range $[0, 1]$	119
6.12 Localization results using Proximity-based classification. Average error = 2.34 m.	120
6.13 Values of the MUSIC spectrum $P_M(\phi, \theta)$ represented in spherical coordinate system centered on the antenna position. The strongest value defines the estimated direction of the incoming signal.	122
6.14 Localization results using a) DOA Estimation, b) Fingerprinting.	123
6.15 At each step of the SOM- $X_{\theta k}$ algorithm the position of each neighbor in proximity of a directional antenna is updated to match their AOA estimate.	125

6.16	The two localization scenario considered: a) a network with four anchors in the corners is localized using SOM-A; b) The same network is localized using SOM-A $_{\theta_1}$; the four anchors are replaced by a single base station capable of estimating the AOA of the messages transmitted by its neighbors.	127
6.17	Localization results using SOM-A $_4$ (four anchors), SOM-A $_{\theta_1}$ (one directional anchor) and SOM-A $_{\theta_4}$ (four directional anchors).	127
6.18	Localization results using SOM-R (four anchors) and SOM-R $_{\theta_1}$ (single anchor).	129
6.19	Performance of SOM-R $_{\theta_1}$ evaluated with different error affecting the AOA estimates. The black dotted line in plot (b) is the SOM-R's error with four anchors.	129
7.1	a) Probability of “connected nodes” for pairs of nodes placed at distance 2.5 m and 6 m; the same probability are also evaluated when the node positions are shifted by a factor $\Delta = 0.5$ m. b) Approximation of the Fisher information $F_{\text{con}}(d = 5 \text{ m}, d_{\text{th}})$ computed by evaluating the cross entropy $D(f \parallel f_{\Delta})$ between the probability mass functions computed with $d = 5$ m and $d = 5 + \Delta$ m, where $\Delta = \{-2, -1, -0.5, +0.5, 1, 2\}$ m.	133

Chapter 1

Introduction

Localization, geo-location, and location sensing are equivalent terms that refer to the process of computing the physical position of a device [104]. At the present, the most popular localization system is the widespread *Global Positioning System (GPS)*, which is used in a variety of military, industrial and recreational applications. GPS devices determine their position by receiving signals from a constellation of 24 satellites arranged in six orbital planes [77]. When the receiver locks on four or more signals, it first uses *Time Of Arrival (TOA)* techniques to estimate the distance of the transmitting satellites. Then, it applies a *multi-lateration* algorithm to compute its 3D position on the earth's surface and report the result to the user.

Although GPS receivers are popular in many consumer electronic devices, some application domains require different localization approaches. Two relevant cases where the GPS technology is not effective are: 1) *Wireless Sensor Networks (WSNs)*, and 2) *Location-Aware applications* operating indoors.

1.1 Wireless Sensor Networks and Indoor Applications

The WSN is a flexible and scalable paradigm that is drawing increasing attention due to its potential utilization in many civilian and military domains. Designed to work without infrastructures, WSNs exploit inexpensive sensor nodes and multi-hop radio communication to implement large-scale monitoring solutions.

Typical WSN applications include environmental monitoring, asset tracking, surveillance and disaster relief [7]. In all these cases, knowledge of the node positions is required to correctly evaluate the results. For example, in precision agriculture, temperature and moisture values are correlated with position to identify micro-climate zones [157]. Knowing the sensor positions is also critical for locating an intruder vehicle in

a military field [10], as well as guiding a team of firefighters to the location of an emergency [33]. Finally, network services such as geographical routing [78], location-based queries [70] and resource directories [98] rely on knowledge of the node coordinates.

The sensor positions are unknown as a result of ad-hoc deployment or because the sensors are mobile. Applications with nodes scattered from airplanes or ground vehicles are examples of ad-hoc deployment; mobile networks are found when nodes are carried by people [165], or attached to vehicles [97] and animals [105]. While some of these WSNs can exploit GPS-enabled sensors, low-cost deployments and applications that can tolerate approximate positioning are better served by solutions that limit, or avoid altogether, the need of GPS receivers. Finding a substitute to GPS is especially important in applications where the nodes strive to be small, inexpensive and low-power.

Other applications requiring implementations alternative to GPS are those deployed indoors, where the signal reception from the satellites is unreliable. Indoor position awareness is required by numerous applications ranging from in-building navigation to asset and personnel management in large warehouses and hospitals [104]. Indoor positioning is also fundamental to enhance the experience of users interacting with ubiquitous computing systems [67]. Smart spaces are expected to sense the user's position to deliver relevant content, facilitate access to nearby resources, and enforce security policies [43, 65].

1.2 Collaborative Localization

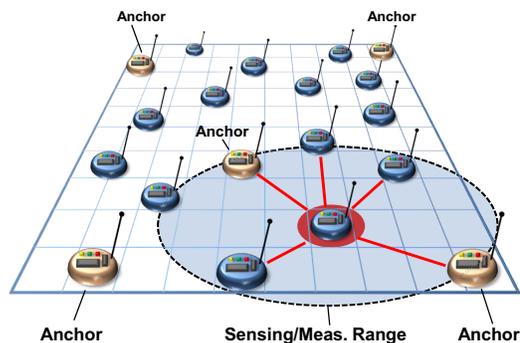
As previously mentioned for the GPS, the absolute position of a wireless device can be determined by collecting measurements from satellites at known position. The same approach can be used to support localization in smaller scale wireless systems: If a node can estimate the distance from three or more *anchor* devices at known locations, its position can be computed using a multi-lateration scheme similar to that used by GPS receivers. This approach is conceptually well defined, but it might fail in applications where nodes have limited communication and sensing capabilities. For example, in low-power sensing applications and other ad-hoc deployments, it might be not possible to guarantee a sufficient anchor coverage to support multi-lateration for every device in the network. In such scenarios, nodes can compensate for the lack of reference devices by taking measurements with other peer nodes at unknown position. If all the nodes participate in this collaborative effort, the collected data can be used to localize the whole network (see Figure 1.1)

Existing collaborative solutions can be grouped in *range-based* and *range-free* schemes depending on the type of measurements used. **Range-based** schemes implement localization using angle or distance

The Collaborative Localization Problem:

Inputs:

- Pairwise measurements between neighbor nodes:
 - Distance → Range-Based Loc.
 - Angles → Range-Based Loc.
 - Proximity → Range-Free Loc.
 - ...
- The position of some anchor nodes - optional: (Anchor-based vs Anchor-free Loc.)



Output:

- Node positions.

Figure 1.1: Collaborative localization in ad-hoc wireless networks. The node positions are computed using information collected by neighboring devices.

estimates between pairs of nodes. This information is typically obtained by augmenting each node with dedicated hardware such as ultrasound transceivers for distance measurements, or directional antennas for angle of arrival estimation. On the other hand, **range-free** schemes only rely on proximity information, i.e. knowing if two nodes are close in space or not. Although this solution can only provide coarse-grained results, proximity constraints are readily available by sensing common physical phenomena such as sound and light, or by exchanging radio messages.

Another classification among collaborative schemes is whether they necessarily assume the presence of anchor nodes or not. While most of the positioning schemes are **anchor-based** solutions that only work if the network contains devices at known position, **anchor-free** schemes will work even when no anchors are present. In the latter case, since no absolute reference points are used, the algorithms generate *relative maps* that are useful to implement services such as navigation, geographical routing [78], and service discovery. A-posteriori conversion into absolute maps is always possible when anchor information becomes available.

Some of the network positioning schemes proposed in the literature are reviewed in Chapter 2. The chapter also provides a formal definition of collaborative localization and discusses some of the reasons that make it a difficult problem: 1) From a theoretical perspective, network localization is analogous to the problem of *embedding* a graph in a Euclidean space. Except for selected cases¹, this problem is NP-hard even assuming error-free range measurements or ideal connectivity [152, 26]. 2) In addition to the computational complexity of finding a solution, the result may be ambiguous, i.e. multiple solutions are admissible, when not enough constraints are available [46]. This situation likely arises in WSNs because nodes have limited

¹The localization problem can be efficiently solved in networks where a high number of error-free inter-node distances and angles are known. See the work using semidefinite programming by Biswas and Ye [22] or the definition of *trilateration* graphs [12].

communication/sensing range and can only interact with a few neighbors. 3) In practice, since nodes use inexpensive sensors and are deployed in uncontrolled environments, the measurements are not only difficult to obtain, but also corrupted by substantial noise that increases the uncertainty in the results [129, 172].

The intrinsic difficulty of the problem explains the large number of localization schemes proposed over the past few years. The literature review in Chapter 2, although far from being exhaustive, serves to define the context for the work presented in the following chapters.

1.3 Outline and Contributions

This dissertation focuses on collaborative **RF-based localization systems**, i.e. systems where the information to localize the network nodes is collected by exchanging radio messages. This approach is inexpensive and available to any node with a built-in wireless interface. In addition, radio message can support both range-based and range-free schemes. In the range-based approach, the *Received Signal Strength* (RSS) measured by the transceiver is used to estimate the distance of the transmitting node. In the range-free approach, the RSS values are usually discarded, but the successful reception of radio messages indicates that two nodes are close in space. This condition is also expressed by saying that the nodes are *neighbors* or *connected*. The proposed work investigates collaborative localization following three research directions:

- The design and validation of novel range-free schemes.
- Theoretical analysis of the limits that bound the localization error for RF-based systems.
- Implementation of practical localization systems.

The following sections describe the contributions in each area and provide an outline of the dissertation.

1.3.1 Range-Free Localization

Chapter 3 describes a novel range-free scheme based on a neural network formalism known as the *Self-Organizing Map* (SOM). The proposed scheme computes the node positions using proximity constraints between sensors and exploits the topological ordering properties of the SOM paradigm (see Section 3). Although other SOM-based localization schemes have been presented, to the best of the author's knowledge, this is the first approach to use SOM for localization based on radio connectivity (see Section 3.2).

Results of extensive simulation show that the SOM-based approach is accurate, computationally feasible, and suitable for a variety of application scenarios. SOM-based localization works with or without anchor nodes. The anchor-free version (SOM-V) generates virtual coordinates that are effective when used for

geographic routing (see Section 3.4). Using the SOM-V's maps, geo-routing achieves performance close to the case where the true node positions are available. A second version of the algorithm (SOM-A) can generate absolute coordinates by including anchor information in the training phase of the map (see Section 3.5). Using only four anchor nodes, SOM-A achieves a localization error as low as $0.3R$, i.e. 30% of the communication range, for medium-sized networks with an average connectivity of just five neighbors per node. This result represents a 30% to 60% improvement over the performance of popular range free schemes such as DV-HOP [127] and MDS [154]. Finally, based on analytical analysis, the proposed scheme is shown to have a low computation and communication overheads, hence making it suitable for resource-constrained networks (see Section 3.6).

1.3.2 The Limits Of Radio-Based Localization

After having presented the SOM-A and SOM-V schemes, this dissertation specifically focuses on the measurements used to implement localization. Despite the attention received by RF-based approaches, two fundamental questions have not been sufficiently investigated in the literature. The two questions are:

1. *What is an optimal approach to infer connectivity information from radio messages?*
2. *How do the performance of the range-free solution compare against those of range-based approaches?*

The first topic addressed is range-free localization. The question raised is how to obtain connectivity information from radio measurements. Since previous research work has often assumed circular connectivity based on an *idealized* radio propagation, only few authors have offered implementation details on how to obtain connectivity information. For example, in the popular *centroid* scheme [30], nodes are connected if at least 90% of the messages are correctly received. Unfortunately, simple heuristics like this one can lead to large errors, especially when most of the nodes are within their radio range.

Section 4.1 adopts a more general connectivity model based on quantization of the RSS values. The problem of finding the optimal quantization level is then investigated on the basis of previous research work that has cast localization as a parameter estimation problem. Using this framework, analysis of the Fisher Information and the Cramér–Rao bound (CRB) serves to identify the optimal threshold that minimizes the theoretical error of range-free localization. The analysis is further extended to avoid computation of the CRB, which requires knowledge of the true node positions. The main contribution of Section 4.2 is the derivation of an approximate formula to compute the optimal threshold and obtain connectivity information from radio measurements. Notably, this approach can be applied to any range-free scheme and avoids the large error typically found when localizing densely deployed networks.

The second question raised is: *How do range-free schemes compare against solutions that use the RSS data (no quantization) to produce distance estimates?* It is known that range-free schemes are capable of *fine-grained* positioning, while the range-free approach can only produce *coarse-grained* results. But range estimates obtained from RSS values can be inaccurate due to the unpredictability of the wireless channel, and some authors have occasionally noticed that connectivity-based approaches can outperform range-based schemes in noisy environments [129, 20]. Given these premises, it is not clear when a system designer should opt for a range-based or a range-free scheme.

Using an approach similar to the one used for the optimal threshold problem, Section 4.3 compares range-based and range-free localization. Analysis of the CRB shows what parameters affect their performance, and under which conditions one approach can potentially outperform the other. Again, the results are analyzed to find design criteria of practical applicability. Section 4.4 shows that an informed choice between the two approaches is possible by comparing the current network connectivity against the value of a function that only depends on the network size and the parameters of the propagation model. While this result and the optimal threshold approach of Section 4.2 are the major contributions of Chapter 4, in general, the proposed analysis serves to understand the factors that affect the localization error in the two cases and suggests strategies to improve RF-based localization.

1.3.3 Implementation of RF-Based Positioning Systems

The last part of this dissertation focuses on practical implementation of localization systems and evaluation of their results in realistic application scenarios.

Chapter 5 investigates a localization scheme that is resilient against different operative conditions. Based on the theoretical analysis of Chapter 4, the SOM algorithm is presented in a new variant (SOM-R) capable of using RSS measurements together with connectivity data (see Section 5.2). The SOM-R scheme retains the quality of range-free localization for low-connected networks, while ensuring accurate localization in dense deployments. Results using RSS traces from three different node deployments show a SOM-R's error independent of the network connectivity and significantly lower than the error of MDS, DV-HOP and the other SOM variants. Similar results are achieved in networks with anisotropic layouts, which are notoriously harder to localize.

Practical implementation of RF-based positioning systems is further investigated by using directional antennas for *Angle of Arrival* (AOA) estimation. Chapter 6 discusses two *switched-beam* antennas developed in collaboration with the Microelectronics Lab, Università Degli Studi di Firenze, Italy. Field tests are used

to evaluate the results of different AOA estimation techniques. In addition, Section 6.2 describes a novel localization system that uses one of the proposed antennas to localize a wireless target using measurements from a single anchor node. Three different localization algorithms are described and evaluated using RSS traces collected during a measurement campaign in a large classroom at the University of Florence. The experiments show that 2D target localization is feasible using a low-cost RF system with a single anchor node. Finally, application of single-anchor localization is extended to collaborative schemes by including AOA information in the training phase of the SOM algorithm. Preliminary simulations confirm the proposed scheme as a viable solution to enable accurate collaborative localization using a single anchor node.

Chapter 2

Background

Localization is an active research area devoted to support *location awareness* in applications where the use of GPS is not cost effective (e.g. sensor networks) or technically feasible (e.g. indoor applications). This chapter presents some of the theoretical background necessary to understand the complexity of the positioning problem. In general, network localization is known to be a computationally intractable problem; additionally, its results are prone to ambiguities. Given these challenges, a large number of localization schemes has been proposed to obtain approximate node positions. Some popular approaches are described in Section 2.3. Additional bibliography is discussed within other parts of this dissertation, and "related work" sections appear at the end of Chapters 3 and 4.

2.1 System Model and Problem Definition

Consider n devices deployed over a two-dimensional space. It can be assumed that each node is identified by a unique ID. Devices communicating using standard protocols, e.g. IEEE 802.x.y, typically have a 48-bit or 64-bit unique MAC address that can be used to identify the node. For simplicity, the unique identifiers are mapped to the first n integers $\{1, \dots, n\}$.

The physical location of each device is described by a coordinate vector $\mathbf{v}_i = [x_i, y_i]^t$, $i = \{1, \dots, n\}$. Applications with nodes deployed in 1D or 3D spaces will consider coordinate vectors $\mathbf{v}_i = [x_i]$ and $\mathbf{v}_i = [x_i, y_i, z_i]^t$, respectively. The goal of a localization service is to compute the unknown vectors \mathbf{v}_i 's using some initial information collected by the devices in the network.

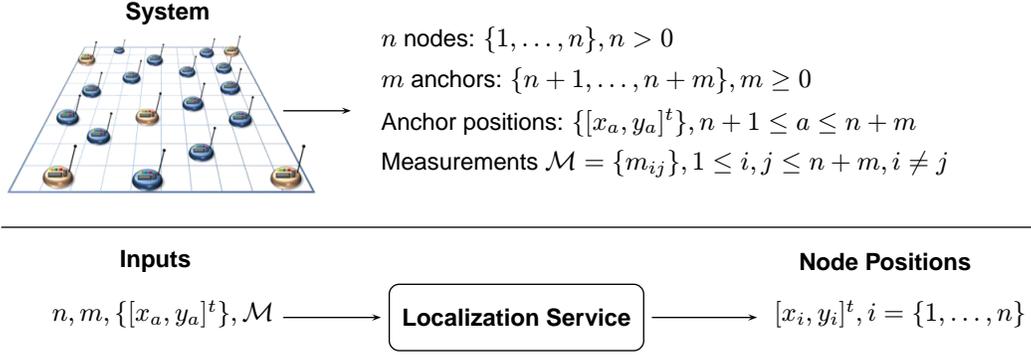


Figure 2.1: A centralized localization system. The unknown node positions are computed using information about the number of nodes n in the system, the position $\{[x_a, y_a]^t\}$ of the anchors (if available), and the set of measurements \mathcal{M} collected between neighboring nodes.

The information available to localize the nodes consists of:

- **A set of anchor nodes (optional).**
- **A set of measurements \mathcal{M} .**

In applications with anchors, some of the node coordinates are known beforehand. If the network contains $m \geq 0$ anchors labeled $n + 1$ through $n + m$, then the vectors $[x_a, y_a]^t, a = \{n + 1, \dots, n + m\}$, are assumed to be known at runtime. This dissertation does not investigate the effect of anchor placement on the localization error. Previous research work has shown that the best results are achieved when the sensors are located inside of the *convex hull* of the anchors [149]. In practice, the anchors should be deployed on the perimeter of the network, preferably on the corners of the deployment. More details on anchor placement and its effect of the localization error can be found in the literature [31, 155, 100].

The set \mathcal{M} contains the measurements available between pairs of nodes in the network:

$$\mathcal{M} = \{m_{ij} : \text{a measurement between node } i \text{ and } j \text{ is available, } 1 \leq i, j \leq n + m, i \neq j\}.$$

Each measurement m_{ij} contains information about the relative position of nodes i and j . For example, m_{ij} can be a binary value that describes the proximity between two nodes, a value that measures their distance, or an estimate of the angular position between two nodes. While it would be desirable to obtain measurements between each pair of nodes in the network, the measurement hardware has a limited sensing range that often restricts the number of data collected. In typical applications, nodes might only be able to collect measurements with a few neighboring nodes located in their proximity. This restriction motivates the use of collaborative localization schemes designed to compute the position of every device in the network,

even those nodes outside the measurement range of the anchors. In contrast, non-collaborative localization schemes requires that each node collect measurements with three or more anchors.

2.1.1 Assumptions

The work in this dissertation is based on the following assumptions:

- **Symmetric Measurements.** It will be assumed that all the measurements are symmetric. For example, if m_{ij} represents the distance between two nodes i and j , the assumption requires that $m_{ij} = m_{ji}$. Asymmetric values will originate if nodes i and j separately estimate the distance to the other node using local measurements. Differences in the measurement hardware and localized interference might results in values $m_{ij} \neq m_{ji}$. By requiring symmetric measurements, the two values will have to be consolidated into a single estimate (e.g. by taking the average of the two values).
- **Centralized Computation.** Another assumption is that the available information can be collected and processed at a central unit. The computational unit can be one of the devices in the network, or a processing device external to the system. The information transmitted to this unit include the number n of devices in the network, the set of measurement \mathcal{M} , and, when available, the number of anchor m and their coordinates $[x_a, y_a]$. It will be assumed that all the devices in the system are working correctly and are capable of reporting their measurements to the central unit. This scenario will be useful to characterize the intrinsic complexity of the problem with knowledge of all the available information. Distributed localization schemes, which have received increasing attention during the past years, will be discussed in Section 3.6 when analyzing the computational complexity of localization using SOM.
- **Static Networks.** The devices are assumed to be static. If nodes move, the measurements in \mathcal{M} are supposed to be collected within a time period during which the node positions do not significantly change. Therefore, a scheme that operates under this condition will approximate the node positions at the measurement time. In mobile networks, the static positions computed at consecutive time steps can be used to implement *target tracking* applications that estimate the trajectory of each node [27, 10]. In such applications, the results can be improved by applying statistical filtering to the data, e.g. [145], or by using the Kalman filter to combine the estimates with a dynamic model that describes the node movement [113]. While the target tracking represent an important problem with numerous practical applications, the work in this dissertation specifically focuses on producing position estimates for static node configurations.

Under the assumptions stated, collaborative localization can be described using two different approaches that use results from graph theory and parameter estimation. The graph theoretical formulation, which is described in the next section, is useful to characterize the computational complexity of the problem. The parameter estimation approach is described in detail in Chapter 4 to analyze the effect of noisy measurements on the position estimates.

2.2 The complexity of the Node Positioning Problem

A wireless network can be modeled as a graph $G = (V, E)$, where the set of vertexes $V = \{1, \dots, n + m\}$ contains an element for each node in the system, and the set E contains an edge $\{i, j\}$ if a measurement between nodes i and j is available (i.e. $m_{ij} \in \mathcal{M}$). Based on the assumption of symmetric measurements, G is a undirected graph.

Given a graph G modelling a wireless network, localization is analogue to the problem of *embedding a graph* in an Euclidean space, a subject that has been extensively studied in the area of computational geometry and graph rigidity. Finding a *graph embedding* consists in determining a mapping function

$$\rho : V \rightarrow \mathbb{R}^d$$

that uses constraints derived from the edge to assigns each vertex to a position in \mathbb{R}^d , where d is the dimensionality of the embedding space. Depending on type of measurements available, different formulations of the problem can be considered.

2.2.1 Embedding With Known Edge Lengths

When some of the inter-node distances are known, the measurements m_{ij} are estimates of the distance between two nodes. The graph embedding problem seeks a mapping ρ compatible with the available data:

$$\|\rho(i) - \rho(j)\| = m_{ij}, \quad \forall \{i, j\} \in E, \quad (2.1)$$

where $\|\cdot\|$ denotes the Euclidean norm. A coordinate assignment produced by the mapping function is called *realization*. The next two sections discuss two relevant problems related to graph embedding.

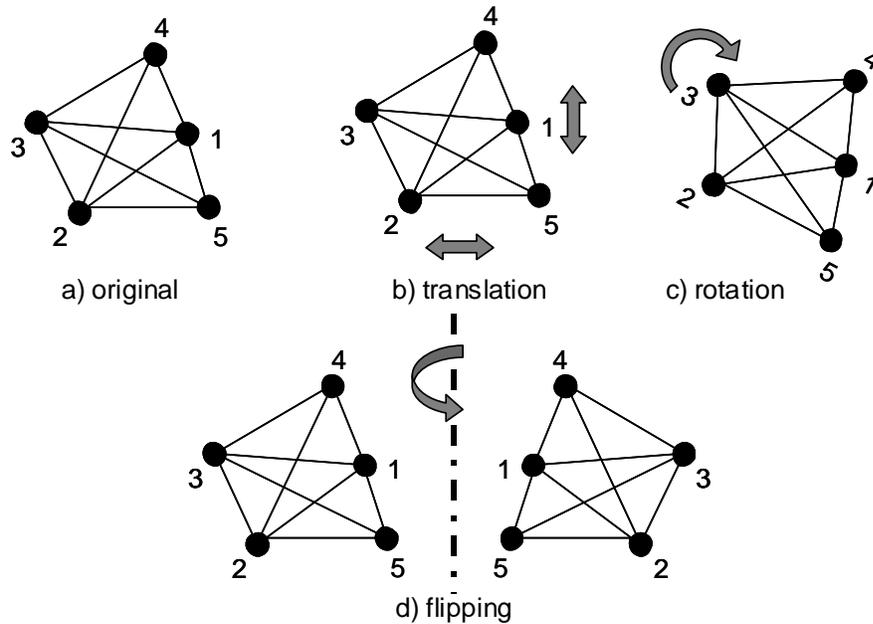


Figure 2.2: Localization ambiguities in absence of reference points.

Conditions For Unique Graph Realization

The first fundamental problem is to determine sufficient conditions for unique realization of the graph G . First, it should be noted that in absence of absolute anchor nodes, every solution will be correct up to global translations, rotations or reflections (see Figure 2.2). The graph can be properly oriented by fixing the positions of three non-collinear nodes in the 2D space, or four such nodes in the 3D space.

Fixing the position of some anchor points, however, does not guarantee an unique solution. When a graph is not *rigid*, the embedding can generate multiple realizations compatible with the available distance constraints. Figure 2.3 shows two graphs where some of the vertices can be moved while maintaining the

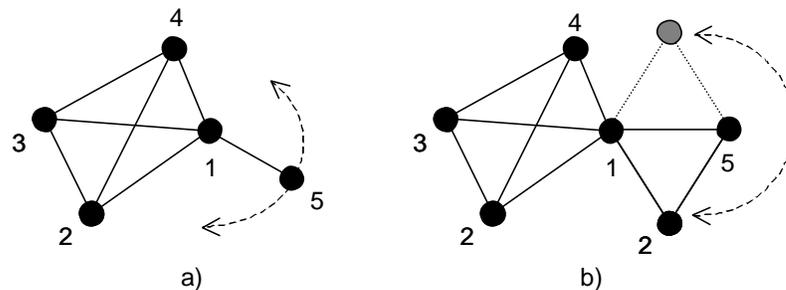


Figure 2.3: Flex Ambiguities: a) Node 5 can be moved continuously along a circular path; b) Node 2 admits two positions compatible with the measured distances from nodes 1 and 5.

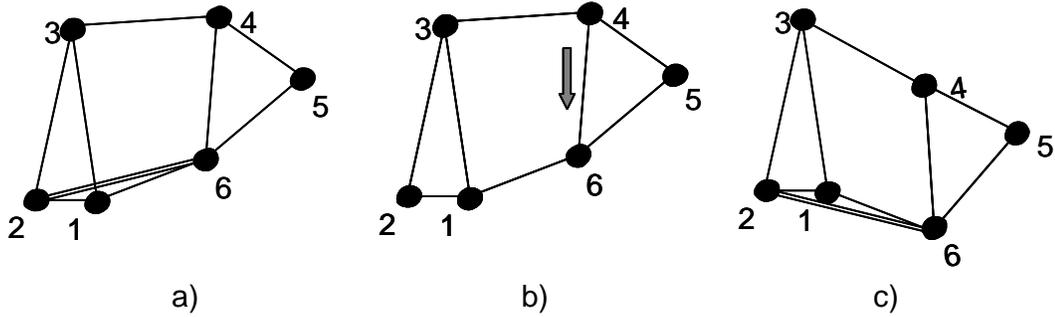


Figure 2.4: Discontinuous flex ambiguities. If edge 2-6 is removed, the graph can be deformed to obtain a new realization where edge 2-6 can be reinserted without changing its length.

same distance from their neighbors. In the 2D space, a graph with n nodes has $2n - 3$ degrees of freedom: two degrees of freedom for each node, minus one for a global rotation, and minus two for translations along the x and y axis. Since each edge introduces a constraint, a rigid graph needs at least $2n - 3$ *well-distributed* edges [92]. But the condition is still not sufficient for unique realization because rigid graphs are susceptible of discontinuous motions. Figure 2.4 shows an example of a rigid graph that can generate multiple realizations without violating the constraints on the edge lengths. If the edge between nodes 2 and 6 is temporarily removed, the quadrilateral defined by nodes 1-3-4-6 can be deformed to generate a new configuration where the edge 2-6 can be reinserted without changing its length. This condition is known as *discontinuous flex ambiguity*.

Unique realization in \mathbb{R}^d requires a graph to be $(d + 1)$ connected and *redundantly rigid*, meaning that the graph is still rigid upon removal of an edge [66]. This condition is necessary and sufficient [71] for unique realization in 2D and can be tested in polynomial time [72]. No similar results exist for graphs in higher dimensions.

The characterization of uniquely localizable graphs is important when localizing ad-hoc networks because nodes have limited sensing range and measurements are only possible with a few nearby nodes. Intuitively, if a network correspond to a graph G with multiple admissible realizations, the localization results produced by any localization scheme are always potentially incorrect. The properties of *rigid graphs* have been used to determine conditions for unique localization of ad-hoc networks [46, 60], or to improve the performance of basic trilateration algorithms under noisy measurements [115].

Most of the graph theoretical results consider scenarios in which no anchor nodes are present. When some of the node positions are known, the distance between these nodes is implicitly known; therefore the graph representing the network should be augmented with an edge for each distinct pair of anchors. In

application with anchor nodes, the condition for unique realizability, i.e $d + 1$ connectivity and global rigidity, should be checked against this extended graphs [12].

Computational Complexity

In addition to possible ambiguities due to an insufficient number of measurements, the graph embedding problem is computationally complex. Saxe has shown that embedding a graph in a Euclidean space is NP-hard [152]. More recently, a number of authors have built on this result to characterize the complexity of the problem specifically for sensor networks, which might contains anchor nodes and where the measurements are likely to be affected by noise. Theoretical results are available for network localization using noisy distance estimates [13, 19]. In all the cases presented the problem is still NP-hard, unless the distance estimates are noise-free and are available for a large-number of nodes [22, 12]. Embedding a graph using local angle information is also NP-hard [28].

2.2.2 Embedding Using Connectivity Information

The results in the previous section apply to problem of embedding a graph with known edge lengths. Such results are relevant to range-based localization scenarios, where nodes posses hardware such as ultrasound or UWB transceivers to measure their distance from the neighbors. When a propagation model is available, the distance can also be estimated using RSS measurements.

A different approach is used by range-free schemes that only rely on *connectivity* information. Several graph theoretical results are also specifically available for proximity-based localization, which is the application considered in Chapter 3 of this work. An overview of the most relevant theoretical results has been presented by O'Dell et al. [131]. In particular, a network with connectivity constraints can be modeled as a *Unit Disk Graph* (UDG), where two nodes are neighbors *iff* their Euclidean distance is less than one. By a proper coordinates scaling, this model can represent an idealized wireless network, where two nodes are neighbors *iff* their distance is less than the communication range R (see Figure 2.5).

Unique Realization of Unit Disk Graphs

Similar to the previous case, the localization problem can be posed as one of embedding an UDG in an Euclidean space. However, the same UDG can be generated by an infinite number of network layouts in which the node positions are perturbed slightly without changing the connectivity between nodes. Therefore, even in presence of UDG admitting an unique realization, the range-free solutions are intrinsically ambiguous.

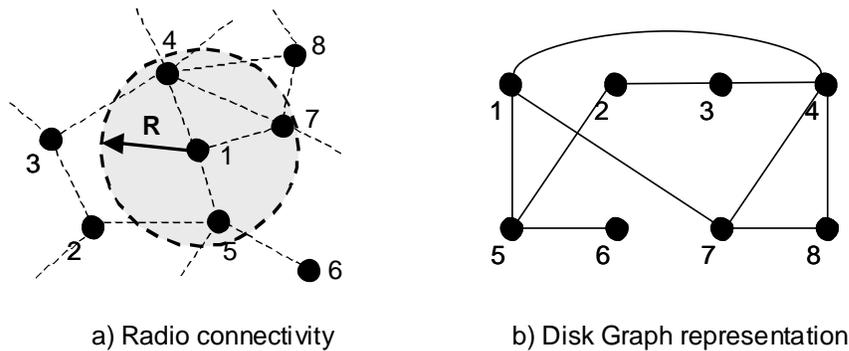


Figure 2.5: A Unit Disk Graph (UDG) used to model a network with connectivity constraints.

Computational Complexity

Embedding an UDG is NP-Complete in one dimension and NP-hard in two dimensions [26]. Recently, the problem has been proved to be APX-hard [108], meaning that the solution cannot even be efficiently approximated. In fact, there exist node configurations for which even an optimal algorithm cannot produce an embedding with quality¹ better than $\sqrt{3/2}$ [89]. The only known algorithm with bounded error has been proposed by Moscibroda et al. [116], who addressed the problem of localization using connectivity constraints (ideal disk connectivity).

2.3 Localization Schemes in Wireless Sensor Networks

The computational complexity of the graph embedding problem and the occurrence of ambiguous realizations help in understanding the challenges faced in designing collaborative localization schemes. Since the node positions can only be approximated, many solutions have been proposed in the literature. Different alternatives are the results of different trade-offs between system complexity and accuracy. For example, reliable estimation of distances and angles requires augmenting each node with dedicated hardware. Therefore, range-based schemes are better suited to high-end applications requiring accurate positioning. On the other hand, range-free approaches trade accuracy for simplicity by relying on proximity information that can be collected using the radio transceiver, or inexpensive RFID tags [68].

The next sections review some popular localization schemes. General survey papers on localization and positioning techniques have been proposed by several authors (e.g. [67, 93, 119, 104]), and several books on WSNs contain chapters on localization (e.g. [179, 141, 20, 160]).

¹The *quality* of an UDG embedding is expressed as the ratio between the maximum distance of neighboring nodes and the minimum distance of disconnected nodes. Ideally, the quality of the realization should be less than one.

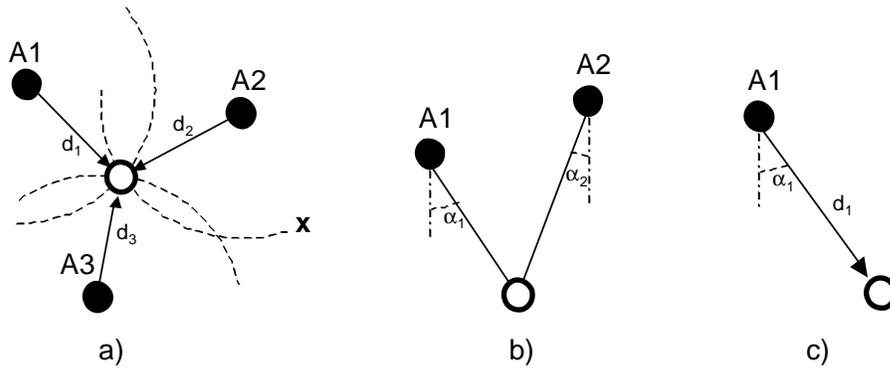


Figure 2.6: Localization in 2D: a) Trilateration: distance measurements from 3 non-collinear anchor points, b) Triangulation: angle measurements from 2 anchor point, c) angle and range from a single anchor point

2.3.1 Range-Based Algorithms

An object in the 2D space can be localized when the following information is known: i) the distance from three non-collinear anchor nodes (trilateration), or ii) the angle from two anchor points (triangulation), or iii) the distance and the angle from a reference point (see Figure 2.6). In the general case, not every node will be able to make measurements with the minimum number of anchors; therefore distance and angle information are often used to implement collaborative localization schemes (see Section 1.2). When a sufficient number of estimates has been collected, the node positions can be computed using multi-lateration algorithms [150], semidefinite programming [21], maximum likelihood estimators [136], or spring-mass relaxation approaches [139]. The next section discusses some popular approaches to obtain range and angle information.

Ranging Using RF Time of Flight (ToF)

RF Time of Flight ranging techniques estimate the distance between two nodes by measuring the time necessary for a radio packet to travel from the source to the destination. Implementation of ToF solutions requires nodes equipped with fast clocks capable of nanosecond accuracy (RF signals travel at 30 cm per nanosecond). In addition, the sender and the receiver must be accurately synchronized.

If synchronization is not feasible with high accuracy, range estimates can be obtained by measuring the *round-trip* time of flight. This solution also requires a precise evaluation of the time used by the target node to process the message and send a reply back to the source. Given these requirements, ToF ranging techniques are better suited to high-end positioning systems requiring high accuracy. Application to WSNs with nodes clocked at only few MHz is not possible.

Ranging Using Time Difference of Arrival (TDoA)

Ranging using acoustic ultrasound is attractive to a wider range of applications mainly because of two reasons: i) ultrasound transceivers are available as COTS components easy to interface with sensor nodes, and, ii) accurate localization can be achieved using low-rate clocks. Given the speed of sound, a 32 KHz clock is sufficient for 1 cm localization accuracy.

Because ultrasound ranging is relatively easy to implement on sensor nodes, several solutions [170, 140, 171, 150, 115] have been proposed where the source node transmits an ultrasound pulse and an RF packet at the same time. The radio message, which travels at higher speed than the acoustic pulse, is used to trigger the receiver node which in turn measures the *Time Difference of Arrival* (TDoA) between the two signals. The distance between the two nodes is computed by taking into account the TDoA and the difference of speed between sound and the RF signal. The main disadvantages of ultrasound ranging techniques is that sound propagation is affected by weather conditions, and the effective range is reduced to only a few meters when the transmitter and the receiver are not aligned (in facts, many transmitters emit a conical directional beam). This limitation, together with the cost of the additional hardware, suggests application to small-scale sensor deployments.

Ranging Using Received Signal Strength (RSS)

The RSS ranging approach is less accurate than ToF and TDoA, but it can be implemented on sensor nodes without specialized hardware. In fact, most of the transceivers used in wireless networks support collection of RSS measurements upon reception of a radio message. The RSS, which measures the signal power received by a wireless device, can be used to estimate the separation distance of the transmitting node. For example, in an ideal free space, the signal decays at a ratio of $1/d^2$; therefore the separation distance between two nodes can be estimated once the transmission power and the RSS are known. In real case applications, the path loss depends on the characteristic of the environment where the communication takes place. For example, the attenuation can be as low as $1/d^{1.5}$ along straight corridors that act as a wave-guide, or it can be proportional to $1/d^4$ for near the ground transmission, where the component reflected by the ground destructively interferes with the LOS (line of sight) component. Additionally, multi-path fading due to reflection, diffraction and scattering of the RF signal causes variations in the received power and ultimately reduces the accuracy of the RSS ranging approach. Despite the unpredictability of the radio signal propagation, ranging using RSS is appealing because it can be implemented in low-cost applications. This approach is discussed in detail in Chapter 4.

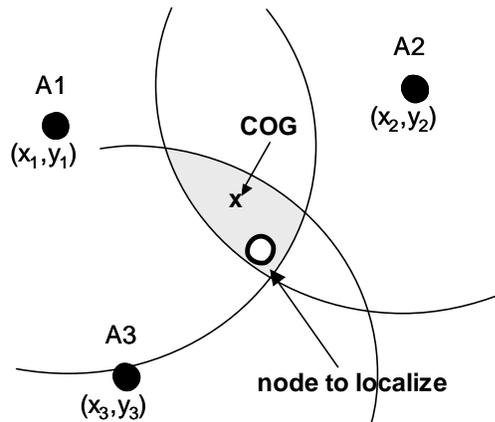


Figure 2.7: Centroid scheme: the position of the node to localize is given by the center of gravity (COG) of anchors heard.

Angle of Arrival (AoA)

Angle of Arrival (AoA) estimation using beamforming or phased antenna arrays has not enjoyed much popularity in sensor network applications due to the cost and complexity of these technologies. However, recent work on directional antennas has demonstrated that simple switched patch units can meet the size and cost constraints of sensor nodes. Several localization algorithms use AoA information to localize sensor nodes [126, 121, 175, 111]. Additionally, base stations equipped with directional antennas can support localization of mobile users in indoor spaces. This topic is covered in detail in Chapter 6. In particular, a switched beam antenna developed in collaboration with the Microelectronics Lab at the University of Florence is used to implement a single-node localization system to track user movements in an indoor space.

2.3.2 Range-Free Algorithms

Range-free algorithms [159] overcome the high cost and system complexity of range-based schemes by using solutions that do not rely on dedicated hardware for distance or angle measurements. The location of each node is estimated by exploiting proximity constraints inferred from radio connectivity or sensor readings. In the first case, nodes that can successfully exchange radio messages are supposed to be within a distance R , where R is the communication range supported by the transceiver. In the second case, sensors can sense natural or artificially generated phenomena that are used as basis for the localization process. The schemes are further classified on whether they rely on the presence of anchor nodes placed at known position or not (anchor-based vs anchor-free).

Centroid

The Centroid [30] scheme is one of most simple, yet popular solutions proposed in the literature. It works by assuming a set of anchor nodes $A = \{a_1, \dots, a_n\}$ placed at known locations $(x_i, y_i)_{i,j=\{1,\dots,n\}}$. The anchors periodically broadcast their coordinates to the nodes at unknown positions. After a sufficient number of messages has been received, each unit determines its location by computing the average value of the anchor coordinates heard. The computed position is the *Center Of Gravity* (COG) of a system of masses placed in correspondence of the anchor nodes heard (see Figure 2.7).

The robustness of the scheme is improved by maintaining statistics on the number of message received from each anchor. Only anchor nodes with a number of successful transmission greater than 90% are used in the computation. The localization accuracy of the centroid method is heavily affected by the number of anchor nodes used. In a subsequent work, the authors propose a solution to adaptively place additional anchor nodes to decrease the localization error [31].

DV-Hop Scheme

In the DV-Hop scheme [127], anchors flood the network with message beacons that are re-transmitted by each node with the hop-count value increased by one unit. Using this approach, each node in the network will eventually be able to compute the shortest path distance (in terms of hop count) from any anchor in the network. To convert the path length into an absolute distance, the average hop count length is computed using the following expression:

$$d_{hop} = \frac{\sum_i \sum_j \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}}{\sum_i \sum_j h_{ij}}.$$

The hop-count distance between any two anchors is used to divide the Euclidean distance separating them (the anchors are at known locations). The result is an average hop count length that can be used to convert an hop-count value into a distance value (see Figure 2.8). Having determined the distance between three or more anchors, each node computes its location using multilateration. The authors use a least square method (the Householder method) to compute the actual position. The scheme works well when the path connecting nodes and anchor nodes lie approximately on a straight line: in this case the hop-count distance is a good approximation of the actual inter-node distance. When the network connectivity is low, or the deployment is anisotropic, the performance degrades since the hop count distance is not a good approximation of the actual distance (see Section 5.3 for more details).

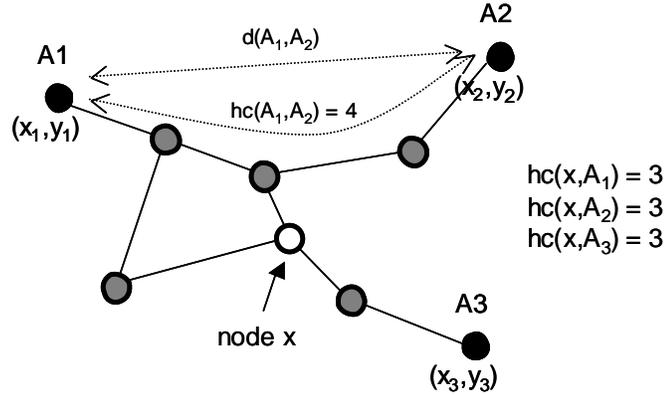


Figure 2.8: DV-Hop scheme: the shortest path (hop-count, hc) is used to estimate the distance from a node x to the anchor nodes in the network. The actual position is computed using multi-lateration.

A similar approach is proposed in [120], but in this case the estimation of the average hop-count length benefits from a priori knowledge of the nodes density through the use of the well known Kleinrock and Slivester formula [80] to determine the hop size:

$$d_{hop} = r \left(1 + \exp(-n_{local}) - \int_{-1}^{+1} \exp\left(\frac{-n_{local} \arccos(t - t\sqrt{1-t^2})}{\pi}\right) dt \right),$$

where r is the average communication range and n_{local} is the local node density.

A.P.I.T.

The APIT scheme proposed in [64] is based on an approximate test to determine if a node is within the triangular area defined by three anchor nodes. If a node were able to move, it would detect increasing (decreasing) RSS levels as it get closer (farther) to an anchor node. The PIT (point in triangle test) determines if a point is inside a triangle by checking for the existence of a direction that would bring the node closer to all of the three anchor nodes (see Figure 2.9a). If such direction does not exist, the node is considered to be inside the triangle. In static deployments nodes do not move; therefore an approximate version of the test (Approximated PIT) is performed by simulating virtual movements in the direction of the neighboring nodes (e.g. by comparing the RSS values seen by adjacent nodes). The final node position is computed by intersecting the area of all the triangles a node belong to and then computing the COG of such area (see Figure 2.9b).

The authors of the the APIT scheme, which is not a truly range-free solution because based on RSS comparisons, report extensive simulation results and comparison with the Centroid and DV-Hop schemes

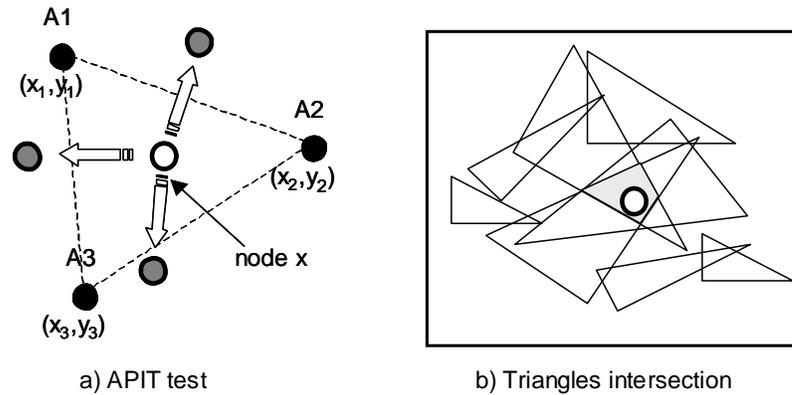


Figure 2.9: APIT scheme: a) Each node uses the APIT test to determine if it is inside the triangle area defined by three anchors. b) The final position is the COG of the intersection of all triangles a node belong to.

presented before. Simulation results show that all the schemes previously mentioned perform well only when a high number of anchor nodes is present and the network density is high. For uniform topologies with connectivity equal to 8, each node needs to receive beacon messages from more than twelve anchors to reduce the localization error under 1.0 R (see Figure 2.10).

SeRLoc

SeRLoc [94] also implements an area-based, range-free approach similar to Centroids and APIT. The anchors are equipped with switched-beam directional antennas capable of covering the 360° horizon with non-overlapping sectors. The antennas transmit *sectorized* beacons to the network nodes along with angular information about the beam used. Similarly to the other approaches considered so far, nodes compute their position by determining the intersection of the beams seen from each anchor node (see Figure 2.11).

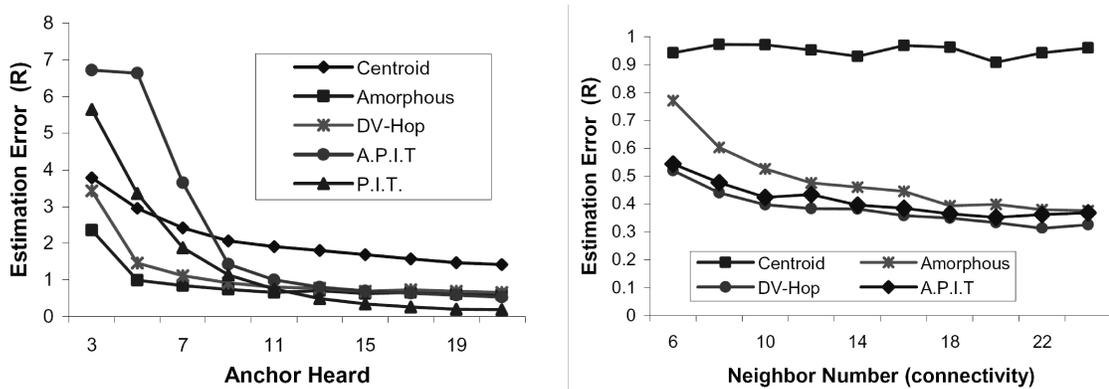


Figure 2.10: Localization error of different range-free schemes reported by He et al. [64].

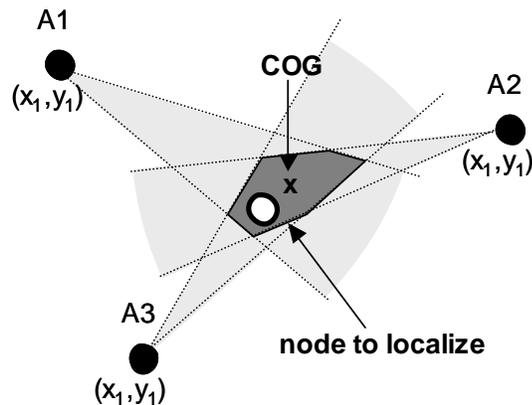


Figure 2.11: SeRLoc scheme: the anchor nodes generate *sectorized* beacons using directional antennas. The final node position is the COG of intersection area.

SeRLoc also addresses the problem of security in sensor network localization. Other algorithms designed to implement secure localization services are presented in [123, 101, 103, 95, 32].

Probability Grid

Probability grid [161] is a localization scheme based on the assumption that nodes are placed on a regular grid. It uses a similar idea to the DV Hop positioning algorithm since anchor nodes flood the network with messages containing their position. Each node estimates the shortest path from each anchor node and then computes the probability of being on each intersection point of the grid. The location with maximum probability is chosen as an estimate for the node's position.

MDS

Multi-Dimensional Scaling (MDS) [25] is a technique that has been extensively used in psychometrics and many other applications to visualize multidimensional data sets. MDS implements a projection technique (to a 2D or 3D space) capable of preserving the similarities present in the original data set. The use of MDS to solve the localization problem in WSN was originally proposed in [154]. The node positions in the 2D space are computed by first creating a $n \times n$ matrix containing the squared distances between each node in the network. If the node distances are not available, the matrix is generated using the hop-count value between each node. The final coordinates are obtained by first double-centering the distance matrix and then using singular value decomposition and retaining the largest two eigenvectors (three for 3D localization). The method has been successively extended to work in a distributed fashion [153, 74, 40, 167], motivated in part by the scarce performance with anisotropic layouts like the ones described in section 5.3.

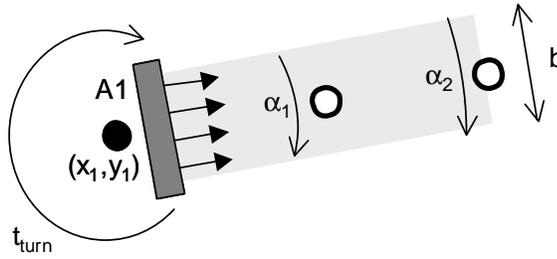


Figure 2.12: Lighthouse: a parallel beam is generated by a rotating base station. Nodes determine the distance from the base station by timing the light beam.

LightHouse

The LightHouse [144] approach exploits the sensing capabilities of nodes. A base station mounted on a rotating support propagates a beam of light having width b (the beam is generated using an array of laser diodes) that is detected by the light sensors mounted on each of the nodes (see Figure 2.12).

Each node computes the distance d from the base station by measuring the time (t_{beam}) during which it sees the light beam:

$$d = \frac{b}{2 \sin(\alpha_1/2)} = \frac{b}{2 \sin(\pi t_{beam}/t_{turn})},$$

where b is the width of the beam, and t_{turn} is the rotation period. Localization in the 2D space is achieved by using three base stations mounted on orthogonal directions.

SpotLight

A similar approach to Lighthouse is used by the SpotLight system [158], which also relies on synchronized light events to localize a set of nodes. Three different scenarios are analyzed (see Figure 2.13): 1) Point scan: if the nodes lie on a straight line (e.g. nodes deployed along a road), they can be localized by a base station that emits a beam of light that is moved at constant speed along the line where the nodes lie. Since each sensor will detect the light at a different time, the node distance from a reference point can be computed by measuring the detection time and dividing it by the beam speed. 2) Line scan: some devices (e.g. lasers) can generate lines of light that can be used to localize nodes on a 2D plane. A first line scanning in one direction (e.g. vertical) allows the nodes to measure their distance from the vertical origin of the deployment area (again, the distance is inferred by the time a node detects the light beam). 2D localization is achieved by a second beam that scans the network in direction perpendicular to the first one (e.g. horizontal). 3) Finally, the third method uses a video projector to illuminate the whole deployment area, which is partitioned in non-

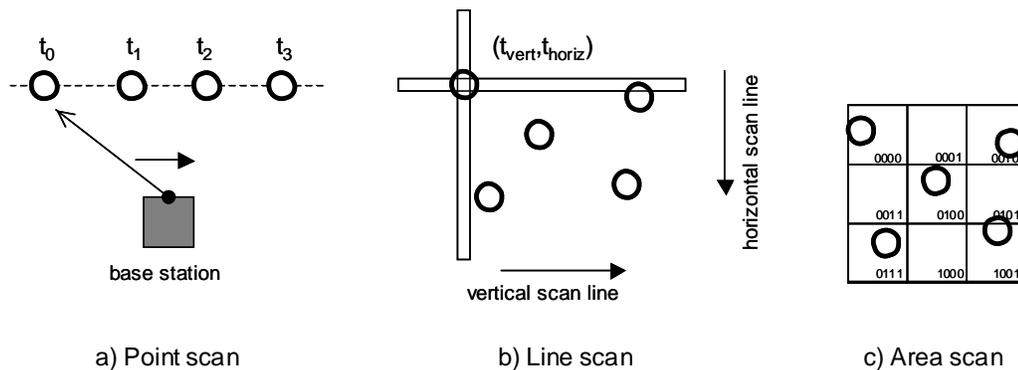


Figure 2.13: SpotLight: localization is implemented by timing the arrive of a light beam. Three options are available: Point Scan, Line Scan, and Area Scan.

overlapping zones. Each zone is illuminated with an unique pattern light, where the presence of light denote a “1” bit and dark is “0”. The temporal sequence of light/dark event is used to transmit a code representative of each area in the network.

2.3.3 Scene Analysis Algorithms

The major problem in using the RSS signal to estimate a distance is that the signal propagation is affected random phenomena such as multipath fading, shadowing, scattering and interferences from other source operating in the same band. The problem is especially severe in indoor environments, where the presence of obstacles (e.g. large metal cabinets, doors, windows, ceiling fans) increases the variability of the signal strength. A localization approach that tries to overcome these difficulties is based on generating RF maps of an environment and then using these maps to locate moving people or objects [14, 106]. The method requires a setup phase during which a mobile device is used to record the signal strength from several base stations present in the network. The result of this phase is the creation of a *RF fingerprints* database that is later used to locate people or objects within the mapped environment. Scene analysis methods have the advantage that can adapt to complex environments and provide acceptable performances when a sufficient number of base stations is available. These methods are also computationally inexpensive, but, on the other hand, they need a time consuming setup phase that needs to be repeated every time there are substantial changes in the environment (e.g. new base stations are added, large piece of furniture are moved, etc.).

Chapter 3

Range-Free Localization Using Self-Organizing Maps

Introduced in the early 80's, the *Self-Organizing Map* (SOM) [82, 83] is a neural network where each neuron contains a weight vector that is updated during the training phase of the map. The neurons are arranged in regular geometric structures, typically two-dimensional lattices with rectangular or hexagonal patterns like the one in Figure 3.1a.

As shown in the following sections, the structure of the map and the learning algorithm result in a versatile architecture that has found numerous applications in the context of exploratory data analysis, pattern recognition and vector quantization. An extensive bibliography of SOM papers has been initially compiled by Kaski et al. [79] and successively updated by Oja et al. [132].

This chapter uses the SOM technique to implement a simple and elegant solution to the range-free localization problem in ad-hoc wireless networks. The use of SOM for node localization is first described in details; then it is evaluated using extensive simulations and comparison with some popular range-free schemes.

3.1 The SOM Learning Algorithm

SOM implements *unsupervised* learning, meaning that the map is able to learn the properties of the training set without the aid of labeled samples or reward functions. Assuming that the information to learn is contained in a large and potentially continuous input set with elements $\mathbf{x}_i \in \mathbb{R}^d$, the map produces a compact

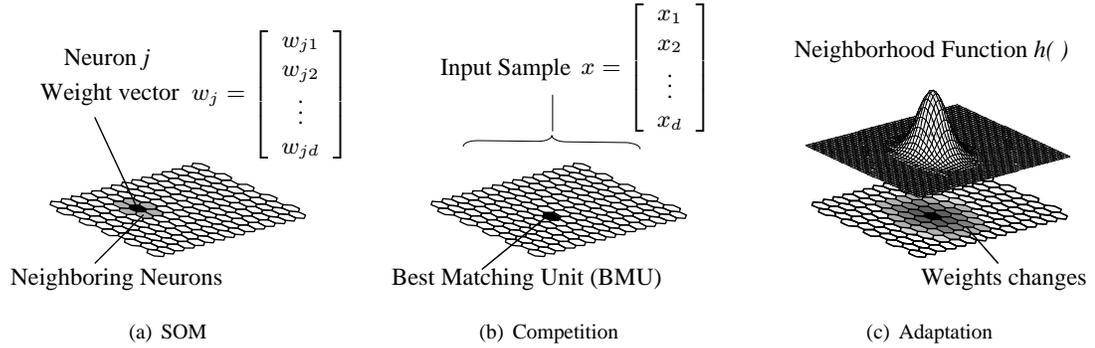


Figure 3.1: a) Self-Organizing Map with hexagonal pattern; b, c) two steps of the training algorithm.

representation of the training set using a finite number of reference vectors $\mathbf{w}_j \in \mathbb{R}^d$. The weights \mathbf{w}_j , also called *models* or *reference* vectors, are initialized with random values and updated by executing multiple iterations of the following three-step sequence:

1. **Sampling:** A sample is extracted from the training set and presented to the network. Let $\mathbf{x}(n)$ denote the sample at the current iteration.
2. **Competition:** The sample $\mathbf{x}(n)$ is compared with the map weights using a distance function. The neuron whose weight is closer to $\mathbf{x}(n)$ wins the competition and become the *Best Matching Unit* (BMU) (Figure 3.1b). If the distance function is implemented using the Euclidean distance, the election rule is:

$$c = \arg \min_j \|\mathbf{x}(n) - \mathbf{w}_j(n)\|, \quad (3.1)$$

where c denotes the index of the BMU, and $\|\cdot\|$ is the Euclidean norm.

3. **Adaptation:** The weight vector \mathbf{w}_c associated with the BMU is updated to increase its similarity with the input sample. During the adaptation process, the BMU activates nearby neurons (*cooperative learning*) allowing them to learn some of the information contained in $\mathbf{x}(n)$. The update rule is:

$$\mathbf{w}_j(n+1) = \mathbf{w}_j(n) + \eta h_{cj}[\mathbf{x}(n) - \mathbf{w}_j(n)], \quad (3.2)$$

where η is the global *learning rate* parameter and h_{cj} is the value of *neighborhood function* that controls the adaptation for neurons close to the BMU (see Figure 3.1c).

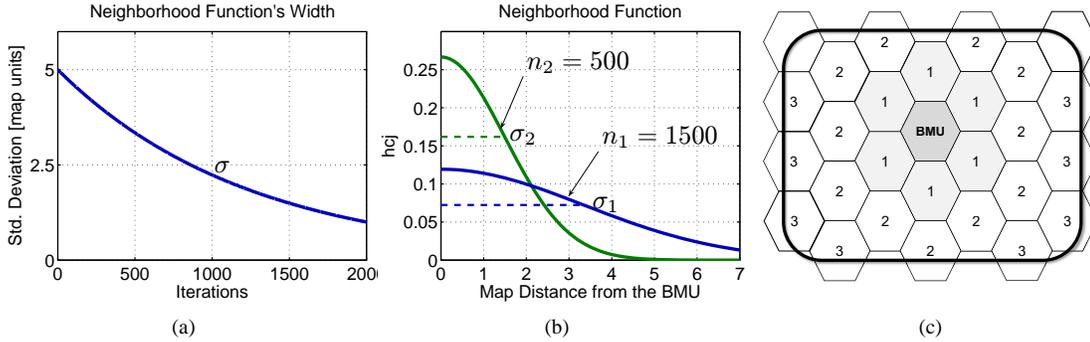


Figure 3.2: a) Exponential scheduling for the parameter σ ; b) neighborhood function at different number iterations; c) portion of a SOM showing the map distances from the BMU.

3.1.1 Learning Parameters

For ensuring convergence, the learning rate η should be computed using a function $\eta(n)$ that decreases monotonically with the number of iterations. Optimal choices of $\eta(n)$ have been discussed in the literature [118, 84]; however, in practice, the exact form of $\eta(n)$ is not a critical factor in the SOM technique.

The update rule (3.2) is also controlled by the value h_{cj} that determines the amount of information learned by neurons close to the BMU. The value h_{cj} can be constant for all the neurons within a given distance from BMU when a *step* neighborhood function is used, or it can be computed using a *smoothing kernel*. As shown in Figure 3.1c, a common choice is to calculate h_{cj} by using a Gaussian function:

$$h_{cj} = \exp\left(-\frac{d_{\text{map}}(c, j)^2}{2\sigma^2}\right), \quad (3.3)$$

where $d_{\text{map}}(\cdot, \cdot)$ measures the distance on the map between two neurons, and the parameter σ controls the *width* of the smoothing kernel. Similar to the learning rate, the parameter σ should be computed using a function $\sigma(n)$ that decreases monotonically. Large values of σ during the initial iterations result in a wide neighborhood function that allows the map to quickly organize the neurons, while the smaller values at the end of the training determine localized changes, allowing the map to describe different input features. Figures 3.2a,b show an example of exponential scheduling for the parameter σ and the resulting neighborhood function at iteration $n_1 = 500$ and $n_2 = 1500$. Figure 3.2c shows part of a hexagonal SOM with labels indicating the distances between the BMU and nearby neurons. Whenever the neurons are updated, the map distance and the current value of σ are used to compute the value h_{cj} that controls the adaptation level of each weight.

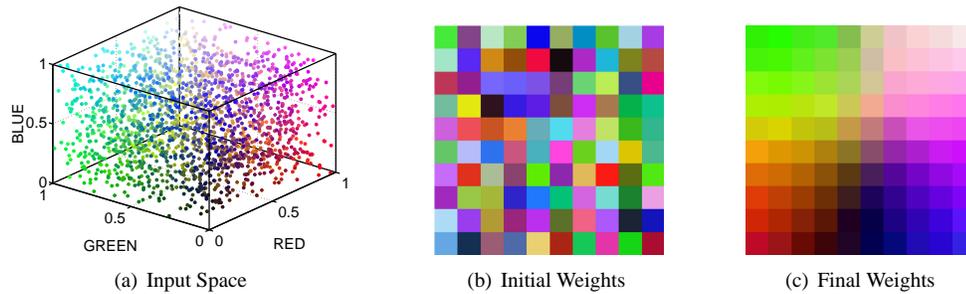


Figure 3.3: 10×10 SOM trained with samples from the RGB color space.

3.1.2 Properties

Some of the SOM algorithm's properties can be illustrated with a simple example. Assume the samples from the RGB color space in Figure 3.3a are used to train a 10×10 map. Both the samples and the map weights are represented by vectors $[r_i, g_i, b_i]$ containing the red, blue and green color components. Figure 3.3b shows the initial SOM with weights randomly assigned. After training the map with a few thousands samples from the input space, the final weights assume the values shown in Figure 3.3c. The results illustrate the following properties:

1. SOM implements a *projection* technique: the three-dimensional input space is mapped onto a two-dimensional surface.
2. SOM implements a *Vector Quantization* (VQ) technique. In this case, 100 vectors were selected as representative values of a much larger input set.
3. SOM generates *topologically ordered* results, in the sense that similar information is mapped to nearby locations. This property emerges as a consequence of the update rule: since adjacent neurons are subjected to similar weight changes, they eventually converge to similar values.

3.2 Localization using SOMs

As seen in the previous section, a SOM can be used to process a large amount of multi-dimensional information and represent it using a compact, low-dimensional model. After training a map, the same election rule (3.1) discussed in Section 3.1 can be used to translate new samples into their corresponding *codebooks* (vector quantization) or to project points onto a two-dimensional surface.

These properties has been used in the past to implement localization schemes for mobile robots in unknown environments [73, 51]. As the robot explores a new space, the multisensory data collected by on-board sensors are fed to a SOM that organizes them on the basis of their similitude. Assuming that sensor readings are correlated with their positions, the SOM defines a virtual map for the space just explored: the robot's location is given by the BMU that matches the current sensor readings. Ertin and Priddy [47] have applied the same concept to the localization problem in WSNs. In their work, synchronous snapshots gathered from the sensors are used to train a SOM, producing a set of weights that define a grid of so-called *virtual sensors*. The node coordinates are approximated by the grid position of the virtual sensor that matches the actual sensor measurements. The authors suggest possible application to target tracking. A similar approach has been used by Sakurai et al. [146] for human tracking in an indoor space, and by Xu et al. [173] for localization of mobile users using RSS from cellular base stations. Finally, Takizawa et al. [163] have proposed a range-based scheme based on an update rule similar to the one used by SOM. These approaches are discussed in more details in Section 3.10.

Proposed Approach

The solution described in this section differs from previous SOM approaches because it does not assume the availability of sensors readings or range estimates, and it does not use the concept of virtual sensors. *The proposed solution uses proximity information derived by radio communication and explicitly compute each node's position during the map's training.* This application of SOM to the node positioning problem is inspired by two simplifying assumptions:

1. The sensor distribution is (approximately) uniform in the deployment area.
2. Nodes that are radio neighbors are relatively close to each other.

Successively, these two assumptions will be relaxed by considering non uniform deployments and more realistic propagation models; however 1) and 2) are useful to illustrate how the SOM technique leads to an intuitive localization scheme.

Imagine that the deployment area is the square region in Figure 3.4a and that a large number of points $[x_i, y_i]^t$ are sampled inside this area and used to train a SOM, say a 5×5 square map. Since the training samples and the map vectors have the same structure, each weight defines a position in the 2D plane. Figure 3.4b shows the values of the random weights, where segments of line are used to link the position of adjacent neurons. As the map is trained, the weights assume the values shown in Figure 3.4c-f. Similarly to

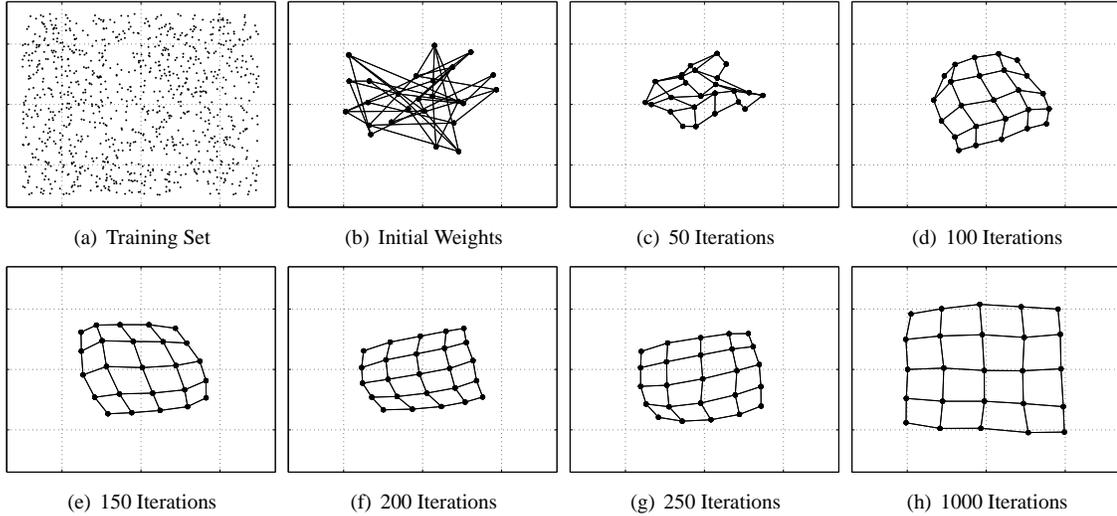


Figure 3.4: 5×5 SOM trained with random samples from a 2D training set.

the example of Figure 3.3, the SOM weights approximate the input distribution, and the weights of neurons that are close on the map converge to similar values. Note that the coordinate assignment in Figure 3.4f is compatible with the positions of an hypothetical 25 node WSN that meets the two assumptions stated at the beginning of this section. More in general, the weights of a SOM trained with points from a 2D uniform distribution can be used as an approximation for the positions of an arbitrary set of wireless nodes. For this purpose, the number of neurons in the SOM needs to match the number of nodes in the WSN, and the map has to be organized in such a way that neighbor nodes are associated to adjacent neurons. The next sections formalize the use of SOM as a tool for sensor network localization.

3.2.1 System Model

Consider a connected network with n nodes placed at unknown locations. None of the nodes is equipped with hardware for position, range or angle estimation, and no assumption is made regarding the availability of sensors; however, nodes can determine their radio neighbors. Let $d_{\text{hop}}(i, j)$ denote the *hop distance*, i.e. the minimum number of transmission required to transfer a message from a node i to a node j .

3.2.2 Modified SOM Model

The unknown node positions are computed using a SOM with n neurons. Each neuron j corresponds to a sensor node and contains a weight vector $\mathbf{w}_j = [x_j, y_j]^t$. This vector, initially picked at random, will eventually contain the estimated location for the corresponding node. The map is trained using the same

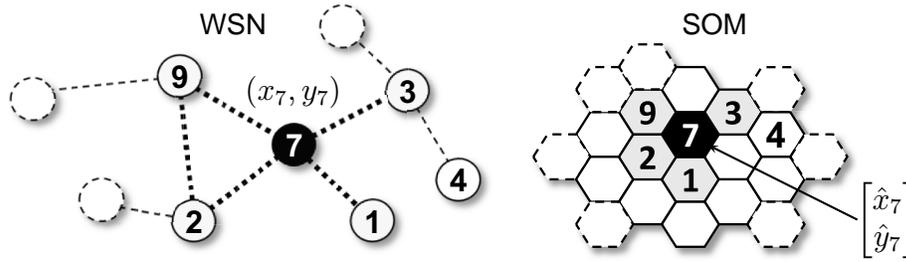


Figure 3.5: Correspondence between the nodes in a WSN (left) and the neurons in a SOM (right). The arrangement of neurons in the SOM describes the neighborhood relationship in the WSN. The weight vector of each neuron contains an estimate of the corresponding node coordinates.

algorithm described in Section 3.1, but with a modified neighborhood function to account for the spatial relationships among the sensor nodes. The new neighborhood function use the hop count distance d_{hop} map distance in place of d_{map} :

$$h_{cj} = \exp\left(-\frac{d_{\text{hop}}(c, j)^2}{2\sigma^2}\right). \quad (3.4)$$

The use of d_{hop} implicitly defines a lattice of neuron with a structure that reflects the hop-count distance between each pair of sensor nodes (see Figure 3.5).

Having described the structure of the weights and the map, the last step involves the choice of a proper training set. To understand how to generate the training samples, note that since no reference points or range information are used, the SOM's results will be correct up to global translations, rotations, flipping or scaling (see Section 2.2). This is a consequence of input used and not of the SOM technique; in other words, any range-free, anchor-free scheme will generate similar results. While these ambiguities might appear as a potential complication, in reality they simplify the algorithm's implementation. In fact, since the result will be expressed in an arbitrarily coordinate system, random samples from an arbitrary distribution (e.g. $0 \leq x, y \leq 1$) can be used to train the SOM.

3.2.3 Localization Algorithm

The algorithm is centralized; therefore each node needs to communicate the list of its radio neighbors to the unit in charge of the computation. Using this information, the hop-count distances between each pair of nodes are computed by first representing the network as a graph, and then applying the Dijkstra or Floyd algorithm. Assume hop distances stored in a matrix D_h with elements $[d_h]_{ij} = d_{\text{hop}}(i, j)$: The matrix D_h is the only input parameter to the algorithm.

Algorithm 1: 2D SOM-V Localization

Input: matrix D_h : hop count distances among nodes

Output: $[x_j, y_j]$ for $j = 1, \dots, N$: node positions

% Parameter Initialization

1: $\eta_{\max} = 0.1;$ $\eta_{\min} = 0.01;$
 2: $\sigma_{\max} = \max_{i,j} \{D_h\} / 2;$ $\sigma_{\min} = 0.001$

3: **for all** nodes n **do**

4: $[x_n, y_n]^T = \mathbf{random}()$

5: **end for**

% Main Loop

6: **for** $n = 1$: to N_ITER **do**

7: $\eta = \eta_{\max} - n(\eta_{\max} - \eta_{\min}) / (N_ITER - 1)$

8: $\sigma = \sigma_{\max} - n(\sigma_{\max} - \sigma_{\min}) / (N_ITER - 1)$

9: $(x, y) = \mathbf{random}()$

10: $c = \arg \min_j \|(x, y) - (x_j, y_j)\|$

11: **for all** network nodes j **do**

12: $h_{cj} = \exp(-D_h(c, j)^2 / 2\sigma^2)$

13: $[x_j, y_j] += \eta h_{cj} ([x, y] - [x_j, y_j])$

14: **end for**

15: **end for**

Algorithm 1 contains the pseudo-code of the localization scheme. The learning parameter $\eta(n)$ and the radius $\sigma(n)$ are scheduled using a linear function that decreases with the number of iterations (see lines 7 and 8). The version of the code described by Algorithm 1 is dubbed SOM-V, because, as discussed in Section 3.4, it generates *virtual coordinates*. Alternative versions (SOM-A, SOM-R) are discussed in Section 3.5 and Section 5.2.

3.3 Simulation Model

The proposed localization schemes have been validated using extensive simulations that were generated in the attempt to model realistic network configurations. Before presenting the results of such experiments, the simulation model is described.

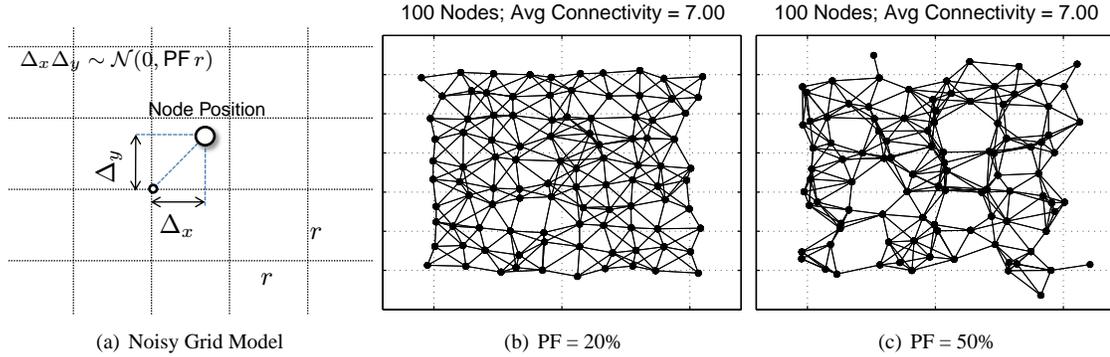


Figure 3.6: a) The noisy grid model; b, c) two 100-node networks with different perturbation factors.

3.3.1 Placement Model

The simulation scenarios are generated according to a *noisy grid* deployment model where the node positions correspond to the intersection points of a regular grid with rows and columns spaced by a factor r (see Figure 3.6a). To capture the random nature of an ad-hoc deployment, each coordinate is perturbed with samples from normal random variables: $\Delta_x, \Delta_y \sim \mathcal{N}(0, \sigma_N)$. The parameter σ_N controls the amount of noise:

$$\sigma_N = \text{PF } r, \quad (3.5)$$

where PF is the *Perturbation Factor* parameter ($\text{PF} \geq 0$) that defines the magnitude of the noise relative to the grid spacing r . Figures 3.6b,c show two 100 node topologies with increasing values of the parameter PF.

Simulations in Section 3.5.4 also consider topologies with node positions sampled from *independent and identically distributed* (i.i.d) random variables. However, the noisy grid model is sometimes more appropriate to describe typical WSN deployments. For example, in most applications such as environmental monitoring and precision agriculture, some control is exerted to ensure an approximate uniform coverage of the monitored area. The noisy grid model also makes it easier to generate connected networks with low connectivity (e.g. 4 or 5), while in the random model, the probability of having connected networks rapidly goes to zero as the communication range is reduced [88].

3.3.2 Connectivity Model

Initially, neighbor nodes are defined on the basis of an *ideal radio model*. If R denotes the maximum communication radius, then two nodes are considered “connected” if their separation distance is less than R , and “disconnected” in the other case. Although this model over-simplifies the nature of wireless communication,

the use of ideal connectivity is intuitive and facilitates comparison with previously published results. A more realistic connectivity model will be considered in Section 4.1.6.

3.3.3 Error Metric

The performance of the proposed schemes is evaluated by computing the average localization error relative to the communication range:

$$\text{Avg. Error (R)} = \frac{1}{R} \sum_{i=1}^n \frac{\sqrt{(\tilde{x}_i - x_i)^2 + (\tilde{y}_i - y_i)^2}}{n}, \quad (3.6)$$

where n is the number of nodes, $(\tilde{x}_i, \tilde{y}_i)$ are the estimated coordinates, (x_i, y_i) are the true node positions, and R is the communication range. Later simulations will use a different error metric to facilitate comparison with the Cramér-Rao bound (Section 4.1.7).

3.4 Anchor-Free Localization: Virtual Coordinates and Geo-Routing

The SOM-V code described in Section 3.2.3 implements the basic version of the SOM localization algorithm. Since only connectivity information are used, SOM-V generates *Virtual Coordinates* [116] that describe the relative location of nodes, in the sense that nodes with similar coordinates are physically close. Virtual coordinates, which facilitate network tasks such as location-based queries and proximity-based service discovery, have found prominent application in the area of geo-routing [78, 90, 142]. By knowing the relative node positions, these schemes achieve efficient packet delivery without the memory overhead of table-driven protocols or the latencies of on-demand approaches.

A direct comparison between virtual coordinates and the ground truth is not possible¹; therefore the performance of a *geo-routing* scheme is used to evaluate the quality of the virtual maps produced by SOM-V. A similar approach has been used to evaluate other range-free schemes, e.g. [125] and [64]. The routing scheme used to validate the localization results implements a simple *greedy* approach: given a source and a destination pair, each intermediate node forwards the message to the neighboring node closest to the destination. The selection rule is:

$$\text{next_hop} = \arg \min_n \|(x_n, y_n) - (x_{\text{dest}}, y_{\text{dest}})\|, \quad (3.7)$$

¹ For range-based localization algorithms that operate without anchor nodes, some quality metrics based on the inter-node distances are available (e.g. see the *Global Energy Ratio* (GER) in [139]). However, in the case of range-free localization, the results are not only possibly rotated or flipped, but also arbitrarily scaled; therefore error metrics bases on the inter-node distance are not applicable.

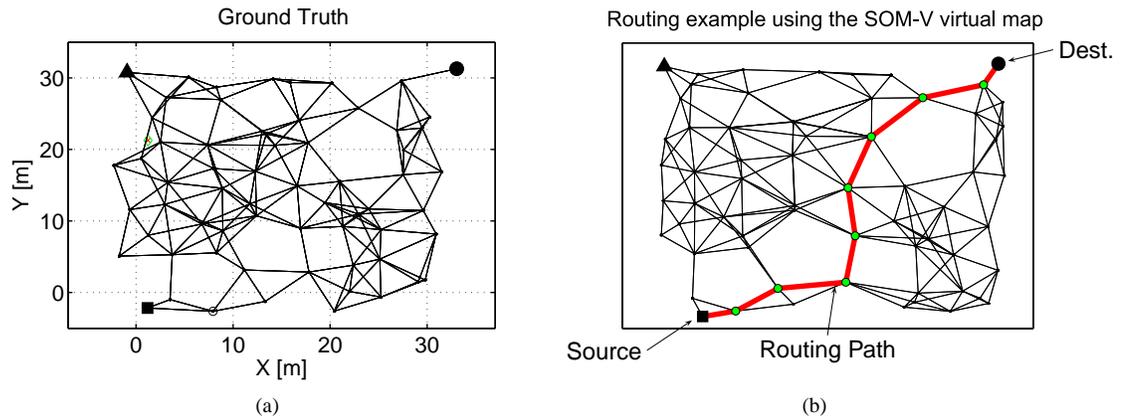


Figure 3.7: a) Ad-hoc network topology; b) Routing path discovered by a geo-routing scheme using the coordinates produced by SOM-V. Note: the two maps has been aligned using the three reference marks on the plot, but this step is not necessary for routing.

where (x_n, y_n) are the virtual coordinates of the neighboring nodes, (x_{dest}, y_{dest}) are those of the destination, and $\|\cdot\|$ denotes the usual Euclidean norm. This basic scheme simply gives up if it is unable to get closer to the destination; however, it defines an useful comparison baseline for more advanced strategies.

Figure 3.7a represents a 64 node network deployed in a square region with side 30 m and PF parameter equal to 25%. The goal is to discover the routing path between two nodes in the corner of the network. This sample application first uses SOM-V to compute a virtual map of the network, and then uses (3.7) to find the route between the two nodes. Figure 3.7b shows the virtual map together with the routing path. The discovered path path is optimal in these in the sense that its length is equal to the minimum hop count distances between the nodes.

A more exhaustive simulation experiment considers 50 topologies similar to the one in Figure 3.7a with PF uniformly selected in the interval between 10% and 50%. For each topology, the rule (3.7) is used to route messages between 50 pairs of randomly selected nodes. Figures 3.8a and 3.8b show the simulation results for different connectivity levels obtained by varying the communication range R . The delivery ratio of the scheme using the SOM-V coordinates is close to the value achieved when using the true node positions, and it rapidly approaches 100% as the connectivity increases. The results show no substantial differences between the lengths of the routing paths produced using SOM-V and the length of those computed using the true coordinates. The two plots are almost completely overlapping in Figure 3.8b.



Figure 3.8: Average performance of a greedy geo-routing algorithm using the true coordinates and the virtual maps produced by SOM-V: a) Delivery Ratio; b) Path Length. The path length plots (right) are almost completely overlapping for the two cases.

3.5 Anchor-Based Localization: Absolute Coordinates

Virtual coordinates are useful to implement efficient packet routing and other network tasks, but some ad-hoc networks require absolute positioning. For example, in a disaster relief application, knowing the sensor positions is necessary to accurately pinpoint the location of an event and provide prompt assistance. To convert relative node positions into absolute coordinates, at least three non-collinear anchor points are needed for the two-dimensional case. When this information is available, the virtual maps are aligned by applying a linear transformation that resolves rotational, scaling and flipping ambiguities. This *a-posteriori* transformation can be used to align the results of any anchor-free localization technique, including SOM-V. The basic SOM algorithm, however, can be modified to include anchors' information in the training phase of the map. This modification not only generates absolute coordinates, but also increases the scheme's accuracy for networks with low connectivity.

3.5.1 Exploiting Anchor Information: The SOM-A Scheme

The *anchored* version of the algorithm, SOM-A, is derived from the basic version by applying three modifications:

1. Weights corresponding to anchors are initialized with the true node positions and never updated.
2. Whenever an anchor node is elected as BMU, the training sample at current iteration is replaced with the anchor's position.

3. The training points are sampled from a distribution whose values are compatible with the deployment area's coordinates.

The first two modifications ensure that weights corresponding to anchors remain in their position. The presence of these fixed points facilitates the map organization during the initial iterations and, assuming three or more anchors, allows SOM-A to generate maps that do not require alignment.

The last modification ensures that weights converge to meaningful values. Different from SOM-V, working with absolute coordinates requires to take into consideration the physical dimensions of the deployment area. In SOM-A, the sampling area is obtained by considering the rectangle enclosing the anchor locations; anchors are assumed to be located near the perimeter of the deployment area, preferably close to the corners. If the network contains m anchor nodes placed at locations $[x_a^{(k)}, y_a^{(k)}]$ for $k = 1, \dots, m$, then, the training points $\mathbf{x}_i = [x_i, y_i]$ are generated by sampling an uniform distribution in the following intervals:

$$\begin{cases} x_i \in [x_a^{(\min)} - \Delta_{ax}; x_a^{(\max)} + \Delta_{ax}] \\ y_i \in [y_a^{(\min)} - \Delta_{ay}; y_a^{(\max)} + \Delta_{ay}], \end{cases} \quad (3.8)$$

where $x_a^{(\min)} = \min\{x_a^{(k)}\}$, $x_a^{(\max)} = \max\{x_a^{(k)}\}$, and the values $y_a^{(\min)}$ and $y_a^{(\max)}$ are computed similarly. The sampling area is expanded in each direction by factor Δ_{ax} and Δ_{ay} to compensate for *border effects* that are notorious in the use of SOM technique. Border effects arise because boundary neurons have fewer neighbors than inner neurons; as a result, the weights along the perimeter of the map are slightly contracted toward the center. To compensate for this effect, the sampling area is slightly expanded by experimentally determined factors Δ_{ax} and Δ_{ay} :

$$\Delta_{ax} = \frac{x_a^{(\max)} - x_a^{(\min)}}{\sqrt{n + m} - 1} \quad \text{and} \quad \Delta_{ay} = \frac{y_a^{(\max)} - y_a^{(\min)}}{\sqrt{n + m} - 1}, \quad (3.9)$$

where $n + m$ is the total number of network nodes (anchors and non-anchors).

3.5.2 Comparison Between SOM-A and SOM-V

Figure 3.9 reports the average localization error for the same set of 50 networks used in the previous section, assuming the presence of three and four anchors in the corners of the map. The results of SOM-V, which are obtained using the *a-posteriori* transformation described earlier, are compared against those of SOM-A. The plots show that the anchored version is more effective in localizing network with low connectivity. For

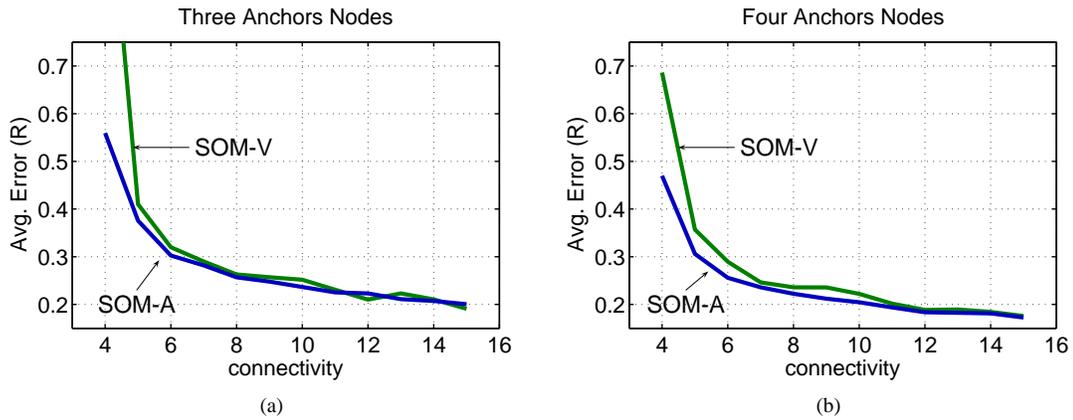


Figure 3.9: Average localization error for a set of 50 random topologies with 64 nodes deployed in a 30 m \times 30 m square region.

networks with connectivity equal to four, SOM-A reduces the error by 52% when three anchors are present, and by 32% when four anchors are available.

Although comparison with other schemes is deferred to Section 3.5.4, the SOM-A's results in Figure 3.9b show a localization error as low as 0.3R for networks with average connectivity equal to five and using only 6.25% of anchors nodes (4 anchors out of 64 nodes). These figures suggest SOM-A as a suitable approach for localization in low-cost deployments with low connectivity and a small percentage of anchor nodes. These networks are likely to be exploited for applications such as environmental monitoring and precision agriculture, where slowly varying signals such as temperature and humidity are monitored over large areas. In these applications, nodes are often placed in sparse configurations to reduce the installation and maintenance costs.

3.5.3 Weight Initialization and Convergence

The results in Figure 3.9 were obtained using random initialization of the weights and 2000 training samples. Existing SOM literature (e.g. [84]) shows that different initialization strategies might influence both the convergence speed and the topological accuracy of the solution. Understanding the effect of the initial weights and number of iterations on the final results is important because the scheme may execute on nodes with limited computational resources.

The previous simulations were repeated by varying the number of iterations and using different initialization strategies. In addition to random initialization, alternative initialization strategies are considered: the AFL and LINE initialization schemes. The AFL scheme has been proposed by Priyantha et al. [139] to

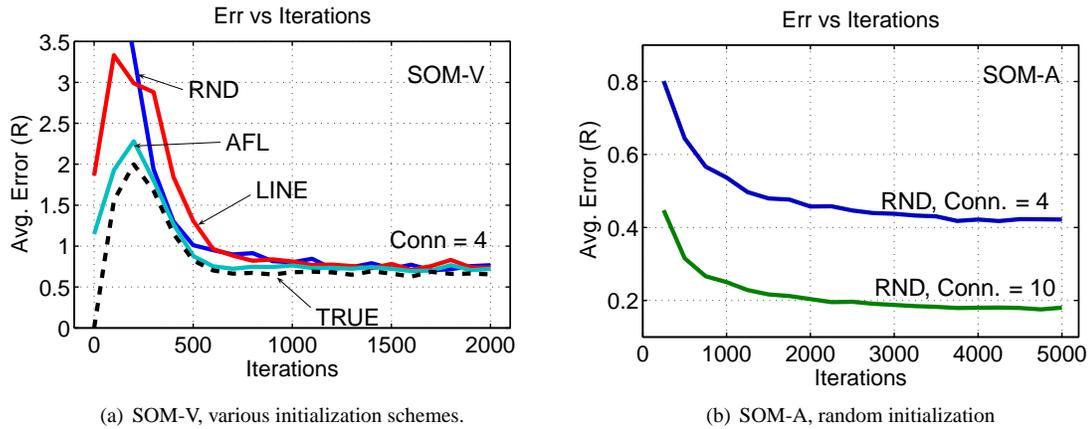


Figure 3.10: Average localization error as a function of the number of iterations used in training the map. The results were generated using the same set of networks discussed in the previous section.

generate *fold-free* initial configurations for a spring-mass based algorithm; the LINE initialization is simple scheme that aligns the initial weights along a line. In a previous work at the IMPACT lab, this heuristic was found to be effective in reducing the occurrence of maps with large topological errors [55]. Finally, a baseline comparison is obtained by considering the error when the weights are initialized with the true node positions.

Figure 3.10a shows the SOM-V’s error as a function of the number of iterations for networks with average connectivity equal to four. This value is considered because preliminary simulations have shown that the differences between alternative initialization schemes are most noticeable for topologies with low connectivity. Even in this case, however, the error plots converge to a similar value as the number of iterations increases over 1000. The weak correlation between the final error and the initial weights is a consequence of training the map with a large neighborhood function. In Algorithm 1, the initial standard deviation for the Gaussian kernel is equal to the radius of the network: $\sigma_{\max} = \max(D_h) / 2$. Such large value causes strong interactions among the neurons; therefore, during the initial iterations, the weight vectors will assume a similar value close to the centroid of the input distribution, regardless of the initial positions. If the neurons were already partially ordered, the convergence speed of the map could be improved by using a smaller value for σ_{\max} to preserve some of the initial information. The work in this dissertation do not pursue this strategy because using a narrower neighborhood function will occasionally result in maps that are only partially ordered (see Figure 3.11).

Figure 3.10b shows the SOM-A’s error for the same set of simulations used in the previous case. When the node positions are computed using SOM-A, the differences between alternative initialization strategies become negligible even for a low number of iterations. Given the minimal differences, the results only

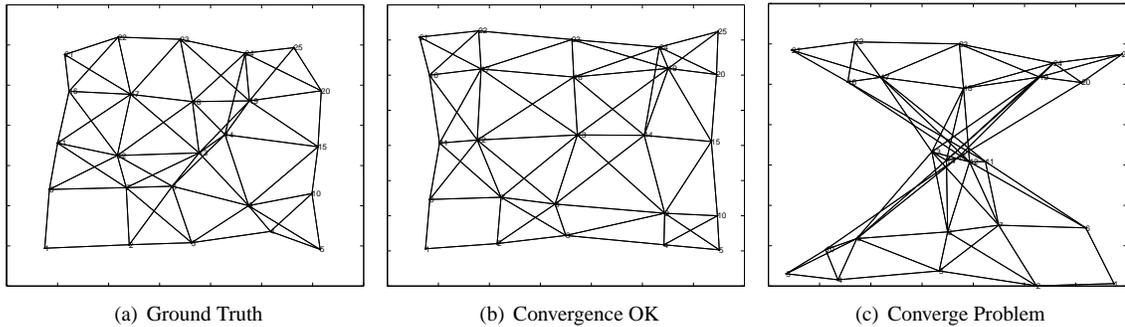


Figure 3.11: Localization example. If the neighborhood function’s initial radius is not wide enough, SOM will occasionally produce incorrect results (case c). This problem is avoided by using the values in Algorithm 1.

show the error obtained using random initialization for networks with low and medium connectivity. The error decreases similarly in both cases and stabilizes around 3000 iteration for connectivity equal to four, and about 2000 iterations when the connectivity is equal to ten. Since the execution time of the algorithm depends on the number of training samples (i.e. the number of iterations), adjusting this parameter provides a mean to optimize the trade-off between accuracy and resources spent (see Section 3.6).

3.5.4 Comparison With MDS and DV-HOP

SOM-A is compared with two popular range-free solutions, the DV-HOP and *Multi Dimensional Scaling*² (MDS) schemes described in Section 2.3.

Figure 3.12 shows the localization results for four sets of 50 random topologies with 64 nodes deployed in a square region with side equal to 30 m. Three of the four sets were generated using the noisy grid model described in Section 3.3.1 with $PF = \{10\%, 25\%, 50\%\}$, while the last set contains networks with node positions sampled from *independent, identically distributed* (i.i.d.) random variables. Note that networks with PF greater than 50% are qualitatively similar to random deployments.

The plots in Figure 3.12 show that the results generated by SOM-A and DV-HOP have a similar trend, but SOM-A consistently produces a lower error. The SOM-A’s results, which were obtained using random initialization and 5000 training samples, are 27% to 45% more accurate than DV-HOP. The differences are more marked for networks with perturbation factor equal to 10%.

In comparing the performance of SOM-A with MDS, the results depend on the network connectivity. The difference are negligible for networks with connectivity greater than ten, but SOM-A significantly outperforms MDS for sparse networks. When the connectivity is equal to four or five, the SOM-A’s error

²The Isomap version that uses the hop-count values as a distance measure between pair of nodes is evaluated [164].

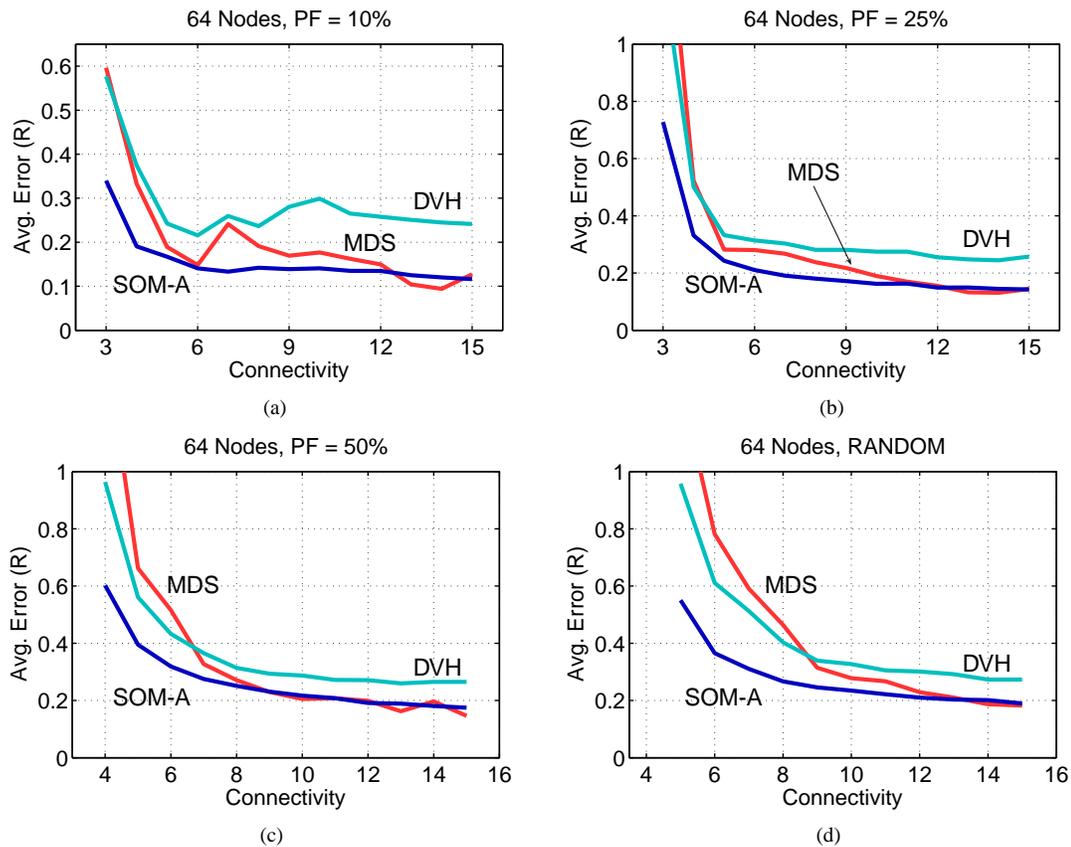


Figure 3.12: Average error achieved by SOM-A, DV-HOP, and MDS in localizing sets of 50 networks with increasing perturbation factors. All networks have four anchors.

is approximately between 40% and 60% lower than the MDS's error. In general, when the connectivity is low, the hop-count distances are poor approximations of the true node distances, especially for nodes that are several hops away. Since MDS equally weights all the available distances, including those with large hop values, the error in sparse networks are usually large. SOM-A does not exhibit this drawback because nodes that are several hops away have a weak effect on each other's positions.

A second set of experiments localizes sets of 25 random networks with increasing numbers of nodes. The connectivity values are fixed to five and ten (see Figure 3.13). In the first case (connectivity = 5), the simulated networks (random deployment) contain up to 200 nodes; given the low connectivity value, above this size it becomes progressively more difficult to generate connected networks. In the second case (connectivity = 10), the networks include up to 350 nodes. The SOM-A technique ensures a localization error around 1.0 R up to 200 nodes and connectivity equal to five. If the connectivity is increased to ten, the error of SOM-A maintains below 0.5 R for networks up to 350 nodes. The results obtained for the other two schemes

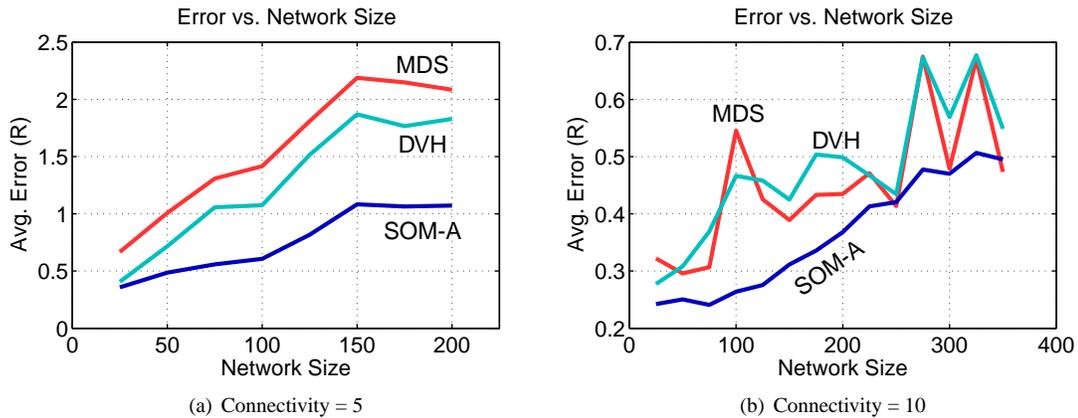


Figure 3.13: Average error of SOM, DV-HOP, and MDS in localizing sets networks with increasing number of nodes and four anchors: a) Network connectivity equal to five; b) Network connectivity equal to ten.

are consistent with those published by Niculescu and Nath [125, 128] for DV-HOP, and those reported by Shang et al [154] for MDS. The error generated by both schemes is significantly higher than the SOM-A's error, especially for networks with low connectivity.

3.6 Computational Complexity Analysis

The SOM-V and SOM-A variants are centralized schemes, but the low communication and computation requirements make them suitable for sensor networks where nodes have limited resources. An analysis of the computational complexity and memory requirements follows.

The SOM algorithms operate on the basis of connectivity information; therefore each sensor needs to communicate the set of its radio neighbors to the unit in charge of the computation. Assuming two-byte node IDs (up to 65536 nodes), the information can be transmitted using a fairly small size radio messages. For example, in a network where the average connectivity is 7, only 14 bytes need to be transmitted by each node. The total traffic can be further reduced by means of in-network data aggregation techniques.

Messages with neighbor sets information are used to generate the adjacency matrix of the undirected network graph requiring $[n(n-1)/2]/8$ bytes of memory space, and then to compute the D_h matrix with the hop count distances between nodes. The solution is obtained by repeating n executions of the popular Dijkstra's algorithm or using the Floyd's scheme. The complexity is $O(n^3)$ in both cases, while the table needs enough storage space for $n(n-1)/2$ elements. The memory requirements for this table can be reduced by taking into account the maximum hop count distance between any two nodes (i.e. the network diameter). The proposed simulations show that most of the 100 node networks with connectivity equal to six have a

diameter lower than 16. Using 4 bits to code the hop-count distances, the size of the table is reduced to $n(n - 1)/4$ bytes of memory. Even if some hop distances were larger than 16, replacing the actual value with the upper limit does not have a noticeable impact on the algorithm because the interactions between units far from each other are weak. Finally, the algorithm needs to reserve the memory space to store the node coordinates (i.e. the SOM weights). Assuming quantized values represented with 2 bytes, the total occupation is $4n$ bytes.

As for the computational complexity of SOM localization approach, the iterative solution allows a trade-off between accuracy and execution time (see Section 3.5.3). Each iteration requires n comparisons to compute the BMU, and the application of the update rule (3.2) to the map weights. Considering that the radius of the neighborhood function shrinks from a value initially equal to the network radius and then goes to zero, the average number of weight updates is approximately $n/2$. Note that, as the width σ of the neighborhood function shrinks, the running time of the solution could be further reduced by only applying the update rule to those weights that are close to the BMU.

Fixed the maximum number of iterations, the SOM training algorithm has linear complexity $O(n)$. However, the simulation results presented in this chapter (see Figure 3.10) and Chapter 5 (see Figure 5.5 at pag. 101) show that increasing the number of iterations for larger number of topologies can reduce the final average error. If the number of iterations is not fixed a priori, but selected as a function of n , the complexity of the algorithm is no longer linear.

Although increasing the number of iterations can improve the quality of the results, using a large number of iterations is not necessary to ensure the convergence of the map. Since the learning parameters are decreased monotonically, the convergence to a stable weight configuration is always ensured.

3.7 Implementation on a Resource-Constrained Sensor Node

The SOM algorithm can be easily implemented in any programming language and it converges to a solution in a limited amount of time. For example, it takes about 0.3 seconds to localize a 100 node network by executing 5000 iterations of MATLAB code on a PC with a 2.66 GHz CPU. More interesting is to evaluate the execution time on embedded hardware commonly used in sensor network applications. The SOM code was implemented using nesC/TinyOS³ and executed on a TelosB [138], a popular COTS sensor node with a 16-bit RISC microcontroller featuring 10KB of RAM, 48KB and working at the frequency of 8 MHz. The

³<http://www.tinyos.net>

N. Nodes	Memory	Exec. time Dijkstra	Exec. time for 1000 iter.
36	0.42 KB	1 sec	62 sec
64	1.48 KB	6 sec	102 sec
100	3.42 KB	22 sec	156 sec

Table 3.1: Memory requirements and execution time of the SOM-A algorithm on a TelosB node equipped with a 8 MHz microcontroller.

algorithm in Section 3.2.3 was slightly modified by replacing the Gaussian neighborhood function with a triangular function. The modified neighborhood function produces similar results using less computation.

Table 3.1 reports the memory occupation of the data structures described above and the execution time to compute the D_h matrix and then to perform 1000 iterations of the localization algorithm. Even using these highly constrained nodes, it only takes about 3 minutes to localize a network with 100 nodes. During the computation, the radio can be turned off and the microcontroller draws only few milliamp of current, with negligible impact on the energy budget of the sensor node. In particular, on a TelosB node, the current drawn by the microcontroller is about ten time less then the current drawn by the radio: the energy spent to localize a 100 node network is about the same energy consumed by the radio in 20 seconds.

3.8 Online vs Batch Training

Algorithm 1 at page 32 implements the standard version of the SOM technique. In the approach used, the map is trained using an online scheme, i.e. the weights are updated at the end of each iteration. Alternatively, the map can be trained using a batch variant in which all the samples are presented to the network before updating the weights [84]. To implement the batch algorithm is sufficient to compute the width σ of the smoothing kernel using a monotonically decreasing function; the use of the global learning parameter η is not necessary. Additionally, the batch version has been shown to yield more stable asymptotic weights than the standard SOM [84, 4].

Despite the potential advantages of the batch training, this method did not improve the error in the positioning application considered in this chapter. In all the preliminary simulations performed, the best results were obtained using the online version of the scheme. The better performance of the standard algorithm is probably explained by the nature of the input set used to train the map. Differently from many SOM applications, the training samples are generated by sampling an arbitrarily large number of data points. In this case, since the data set is not fixed, there is not advantage in computing the weight updates after having presented all the training points to the map.

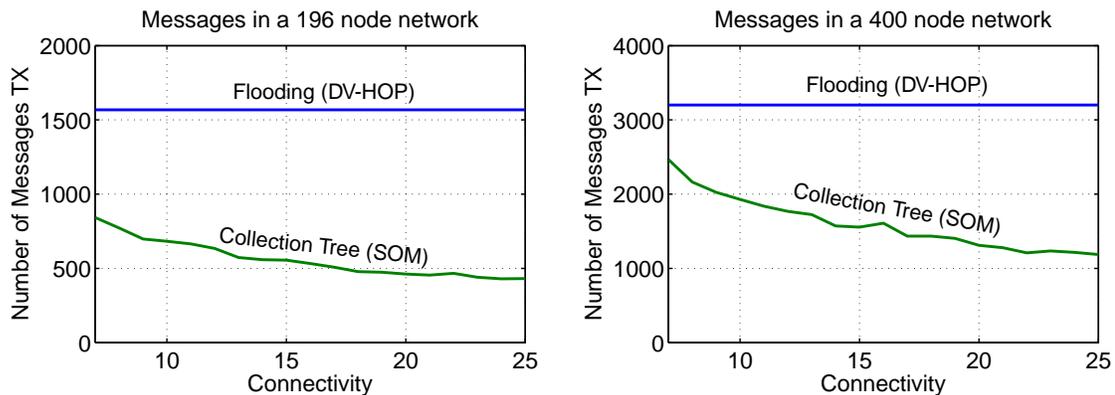


Figure 3.14: Number of messages transmitted in a 400 node sensor network to enable localization with DV-HOP (distributed) and SOM (centralized). Neither flooding nor the collection tree protocol have been optimized.

3.9 Comparison with Distributed Localization Schemes

Recently, several research efforts have been directed toward the study of distributed localization algorithms. This interest is motivated by some limitations of the centralized computation model. Centralized localization is not a viable solution when: 1) The communication overhead to transfer the input data to a central unit is too high; 2) None of the devices in the system possess the computational resources to compute the whole solution; 3) The results are critical and introducing a single point of failure will put the reliability or security of the system in jeopardy; 4) The application require privacy; therefore, similar to GPS, the location should be estimated by the device itself.

In applications requiring privacy, the use of a centralized scheme like SOM will not satisfy the requirements. In other application scenarios, a centralized scheme might be preferable to a distributed algorithm. As shown in the previous section, the complexity of SOM is linear in the number of nodes, and the algorithm can be executed on hardware with limited resources. Given the modest computational requirements, not only the SOM scheme can be executed on a single device, but it can also be deployed on a few back up units to improve the system reliability. Additionally, the majority of the nodes not involved in the computation will only run the application software, thus reducing potential failures due to conflicting software modules.

Another factor to take into consideration when comparing localization schemes is the number of messages transmitted within the network. In most sensing applications, the nodes are pre-programmed to report their readings to a central unit. Low-power tree collection protocols are available as part of the ZigBee standard [86] and the TinyOS [6]; other custom implementations are available from radio chip manufac-

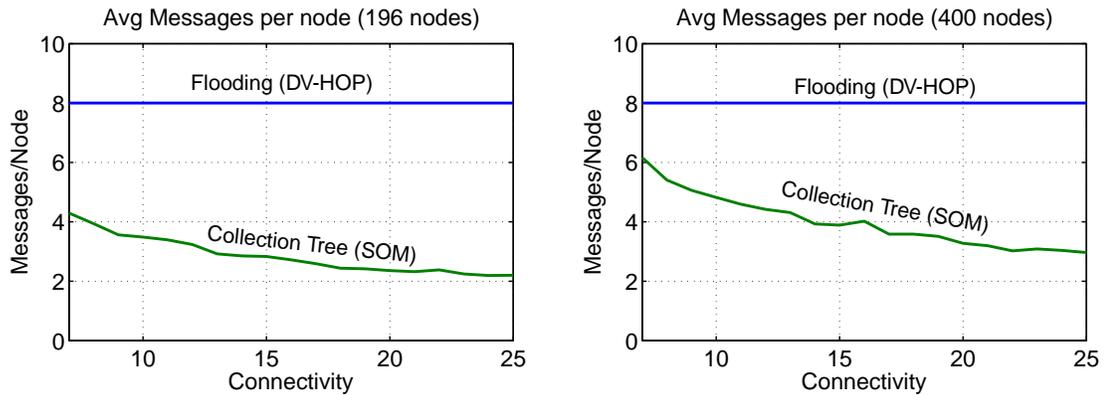


Figure 3.15: Average number of messages per node transmitted in a sensor network to enable localization with DV-HOP (distributed) and SOM (centralized).

turers (e.g. [3]). The information about the neighboring set of each node, i.e. the input data required by SOM, could be *piggybacked* on the sensor readings already transmitted without a significant communication overhead.

Even if the information required by SOM had to be transmitted independently, in some cases the energy requirements will be lower than those of a distributed scheme. Consider the DV-HOP algorithm used for comparison in this chapter. The DV-HOP is a good example of a distributed scheme: Each node computes its own position using information from the network. However, this approach requires each anchor to flood the network with two waves of messages: the first one to compute the hop count distance, and the second to transmit the scale factor for each hop. In a network with four anchors, each node will transmit eight messages. Other popular distributed schemes such as *N-hop multilateration* [151] and *Robust positioning* [148] use a similar approach (see the work of Langendoen and Reijers for a detailed comparison between the three approaches [93]).

Figure 3.14 shows the number of transmissions required to implement SOM and DV-HOP in two randomly deployed networks with four anchor nodes and number of nodes equal to 196 and 400, respectively. For SOM, each node transmits its neighbor list by sending a packet to a sink node that, in the simulation, is supposed to be in the center of the network; therefore the same packet have to be re-transmitted several times until it reaches the sink. Despite no data aggregation techniques are used, the number of messages is significantly lower than a scheme that uses a flooding mechanism such as DV-HOP. The same results are shown in Figure 3.15 by reporting the average number of transmissions per node. For example, on a 400 node network with connectivity equal to 15, each node needs to transmit, on average, about four messages, exactly half of the messages required by DV-HOP. Therefore, if the node positions will be used at the central

unit (e.g. to interpret the sensor data), a centralized solution such as SOM might result in a communication overhead lower than that of a distributed algorithm such as DV-HOP.

If the positions have to be used by the node themselves (e.g. georouting), the energy expenditure to transmit the results back to the network might render a centralized scheme not attractive. In general, while the two approaches will have to be evaluated case by case, the author believes that the use of centralized scheme could be beneficial even in large scale deployments. Other considerations on the two approaches are discussed in Section 5.3.

3.10 Related Work

This section focuses on previous localization research using SOM. The relation between the theoretical aspects of localization and convergence results available for self-organizing maps are also discussed.

3.10.1 Localization Using SOM

Ertin and Priddy [47] have used SOM to solve the localization problem in WSNs. Their model is based on the assumption of devices capable of sensing a common phenomena, such as acoustic or seismic, at synchronized time steps. A further assumption is that the correlation between sensor readings s_i and s_j from nodes i and j is a function only of the distance between nodes: $E[s_i s_j] = f(\|p_i - p_j\|)$, where $p_i = (x_i, y_i)$ and $p_j = (x_j, y_j)$ are the physical location of the two sensors. The input samples used to train the SOM are obtained by concatenating the sensor readings collected at each time step: $\mathbf{x}_n = [s_1^{(n)}, \dots, s_N^{(n)}]$, where $s_i^{(n)}$ is the output of sensor i at time step n . Once the map has been trained with the samples \mathbf{x}_n , each neuron contains a weight vector $\mathbf{w}_j = [w_{j1}, \dots, w_{jN}]$ whose dimensionality is equal to the number of nodes in the network. At this point, there is not a direct correspondence between the neurons and the position of the sensor nodes yet, however, the authors suggest that each sensor i can be associated with the neuron j having the largest component w_{ji} . The relative position of such neuron j in the lattice of neurons defines the virtual coordinates of sensor i . No numerical results are provided to characterize the accuracy of the solution, but the authors qualitatively describe a possible application to the target tracking problem.

A similar approach has been used by Sakurai et al. [146] to implement a tracking application for people moving inside a building. Similarly to the previous case, the input samples used to train the map contain the value sensed by the n sensors installed in the monitored area: $\mathbf{x}_n = [s_1^{(n)}, \dots, s_N^{(n)}]$. In this application the SOM is not used to compute the physical location, but to create a *logical map* where sensor

readings with similar values are grouped together. Again, only a qualitative analysis of the result is presented. Numerical results are instead provided by Xu et al. [173], who have used SOMs to track the movement of people in large outdoor areas using signal strength values measured from nearby cellular stations. In their case the training samples are given by the RSS value collected by the mobile users as they move among the cells covered by several base-stations. This approach and the others described above are substantially different from the scheme presented in this work and have more resemblance with other *fingerprinting* localization techniques (see, for example [14, 107]).

Takizawa et al. [163] have proposed a distributed range-based scheme that uses some of the concepts found in the SOM technique. In this approach, the nodes use a modified version of the update rule (3.2) discussed in Section 3.1 to iteratively update their position. Interaction between nodes is limited to 1-hop and 2-hop neighbors. This method is similar to the refinement phase used in several range-based schemes (see, for example [151, 148]), and it is susceptible to convergence to local minima. An heuristic solution is proposed to avoid this situation. Paladina et al. [133] have also proposed a distributed localization scheme based on the use of SOM. Their model assumes nodes deployed in a regular grid, therefore each node can be thought as positioned in the center of a small 3×3 SOM where the remaining eight neurons contain the position of the surrounding one-hop neighbors. Each node uses this small SOM to process the the positions transmitted by its neighbors and compute its own position, which is then propagated to the remaining nodes.

The SOM approach described in this chapter is analogous to previous applications of the SOM technique to graph drawing [114, 24], a branch of graph theory that deals with the visualization of complex graphs. The graph layout problem is similar to localization in the sense that it also seeks to find a coordinate assignment such that vertices connected by edges are positioned close to each other. But, while the evaluation of a graph layout is mostly based on aesthetic factors (e.g. uniform distribution of nodes and edge lengths, separation between graph elements, number of edge crossing), the results of the localization assignment are directly comparable with the true sensor locations. This work explicitly focused on reducing the localization error of SOM maps.

3.10.2 Theoretical results

The theoretical results discussed in Chapter 2 help in understanding the intrinsic difficulty in computing the node positions and why, in the general case, only approximate solutions are available. The merit of applying the SOM technique to the localization problem is that it provides a low-complexity solution that has been shown to produce accurate localization results in different localization scenarios. Regarding the SOM tech-

nique itself, despite the attention received, self-organizing maps algorithm have proven to be very resistant to mathematical characterization and theoretical results are only available for one-dimensional configurations of neurons. The first formal proof on ordering and convergence properties of SOM has been presented by Cottrell and Fort [41] for uniform distribution of the input samples and a step-neighborhood function. The proof has been extended to more general neighborhood function by Fort and Pages [49], but theoretical results for the two-dimensional case are still incomplete [42].

3.10.3 Discussion

A lack of formal proof in the general case does not necessarily penalize this approach with respect to other techniques. Given the possible ambiguity in the localization results and additional uncertainty caused by the noise in the measurements, even a solution with proven convergence properties would not be guaranteed to converge to the ground truth. At present, simulations and test-field experimentations are the only tool available to compare the performance of different localization schemes working under realistic system configurations. This is also the approach followed in developing the results in this chapter.

Chapter 4

Understanding the Limits of RF-Based Localization

The SOM localization algorithm discussed in the previous chapter implicitly assumes the existence of a service capable of determining whether two nodes should be considered neighbors. This chapter specifically focuses on the measurements available by exchanging radio messages, and, in general, on the performance of RF-based localization. Two fundamental problems are addressed. The first one is how to convert the information collected by the transceiver into connectivity constraints (see Sections 4.1 and 4.2). The second problem is how to decide when to implement a range-free or a range-based scheme (see Sections 4.3 and 4.4). Since radio messages support both approaches, solving these problems has practical implications for implementing RF-based localization systems.

The analysis in the following sections adopts a *parameter estimation* approach based on evaluation of the Fisher Information and the Cramér–Rao bound (CRB). The presented results serve to understand the parameters that affect the performance of the two approaches and suggest strategies to reduce their error.

4.1 Localization Based on Radio Connectivity

As discussed in Chapter 2, radio messages support an inexpensive approach to obtain proximity information. The principle is simple: since each node has a limited communication range, the successful transmission of a radio packet from node A to node B implies that the two nodes are close in space. The use of connectivity information is also appealing for the following reasons:

1. Since nodes already exchange data using radio messages, connectivity information is easy to acquire or it might be already available; in fact, many contention-free MAC protocols and routing algorithms also require this information.
2. Connectivity between a pair of nodes is a binary value. This one bit information can be efficiently communicated across the network with minimal impact on the energy budget of sensor nodes.
3. Several localization schemes are available to process connectivity data on hardware with limited memory and computational resources (e.g. [30, 55]).

Another merit of connectivity-based localization schemes is that they are easy to simulate. Using the *idealized radio model*¹ widely adopted in previous research work, connectivity between nodes can be simulated regardless of the complex phenomena that regulate RF propagation. However, although the idealized radio model provides an abstraction useful in simulation studies, it does not define a criterion to obtain connectivity data in real world applications. In other words, system designers implementing a connectivity-based scheme will have to define their own rule to establish which nodes are to be considered neighbors. In the Centroid scheme [30], for example, nodes are regarded as neighbors if at least 90% of the message transmitted are successfully received. Unfortunately, simple, but arbitrary rules like this one will not always produce satisfactory results.

4.1.1 Motivating Example

Consider the case where one wants to localize the nodes in Figure 4.1a. The data² for this network has been collected by measuring the average RSS between pairs of nodes in the cubicles of an office space [136]. Note that every node of this network is in the radio range of every other node, and no packet loss was reported. Application of a connectivity rule based on percentage of received packets produces a **fully connected network**.

To get a sense of how range-free localization works in this scenario, nodes 3, 10, 35, 44 are used as anchors, while the remaining nodes are localized using three different schemes: DV-HOP [127], Multidimensional Scaling (MDS) [154], and localization using Self-Organizing Maps (SOM) [55] described in the previous chapter. The localization results are reported in Figure 4.1b,c,d. In all of the three cases the position estimates are largely incorrect; the average error is between 4.97 m for SOM to 10.4 m for MDS.

¹Two nodes are connected if their distance is less than a fixed radius.

²<http://www.eecs.umich.edu/~hero/localize/>

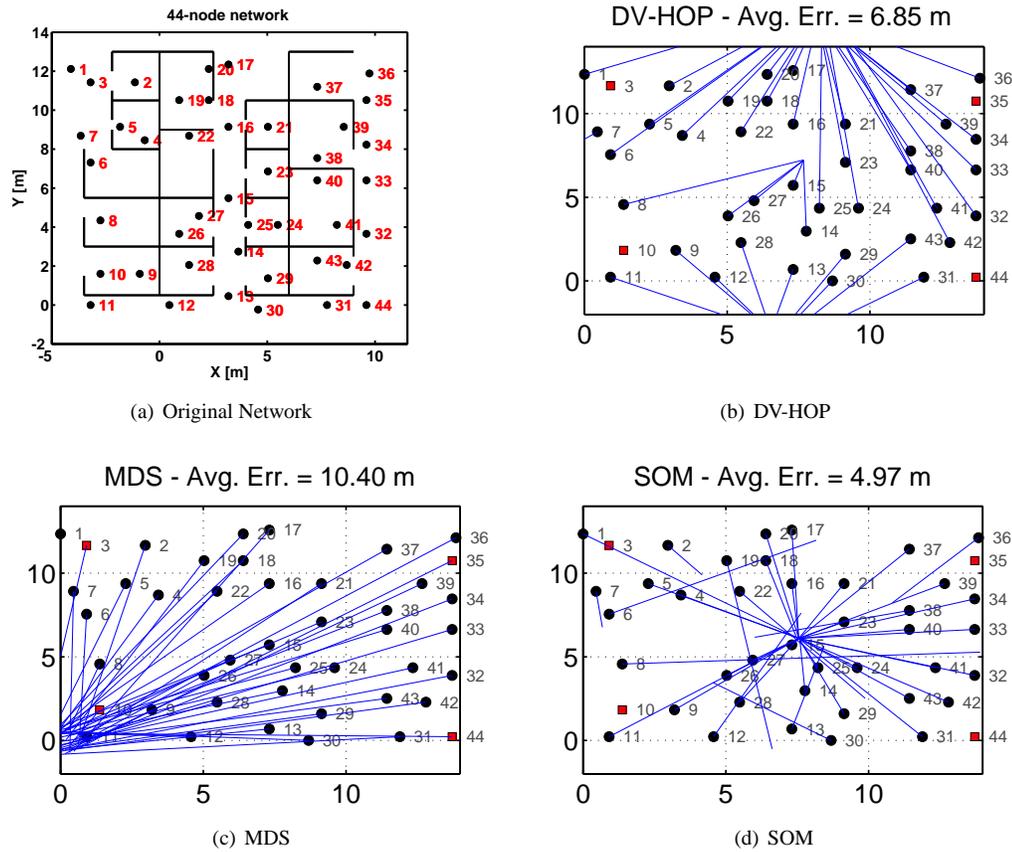


Figure 4.1: Localization errors for a 44-node network deployed in the cubicles of an office space. a) Original Network; b,c,d) Localization results using DV-HOP, MDS, and SOM. Segments of lines are used to connect the true node position to the estimated ones. The long lines in all the three plots denote large errors.

The large localization errors in Figure 4.1 are not surprising. Since the network is fully connected, any connectivity scheme will try to position the nodes close to each other, thus resulting in a large error. *Accurate localization using range-free approaches is not possible in dense networks, because connectivity data carry little information about the node positions.*

To overcome the limitation of range-free schemes in densely deployed networks, one can “artificially” reduce the connectivity by setting a threshold and considering neighbors only those pairs of nodes whose average RSS exceeds the threshold. It is not clear, however, how such a threshold should be set: a value that is too low might be ineffective in reducing the connectivity, while a value that is too high might cause the network to become disconnected and, again, results in large localization error.

What is the correct threshold value? Figure 4.2 provides an empirical answer to this question by reporting the average localization error for different values of the RSS threshold. The plots show that a proper

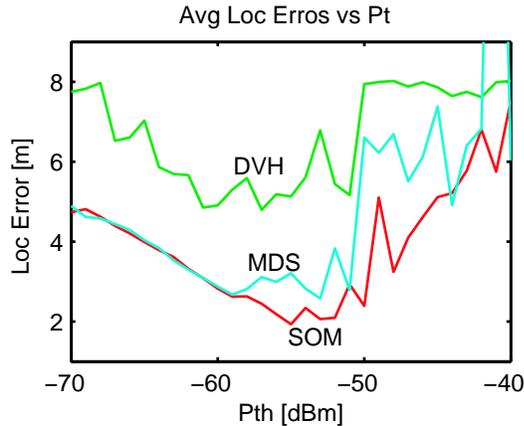


Figure 4.2: Localization error of SOM, MDS, and DV-HOP for the network in Figure 4.1a. The error is plotted as a function of the threshold P_{th} used to quantize the RSS data.

threshold should be between -60 dBm and -50 dBm; in fact, in this range all three algorithms produce a low error. However, the error plots in Figure 4.2 are computed using knowledge of the true node positions. Computing the localization error is possible only if the true node coordinates are known; therefore, in real-world applications, an effective threshold value will have to be found using an alternative approach.

RF-based localization is a popular research topic, but the problem of how to convert RSS measurements into connectivity data has not been thoroughly investigated. The solutions proposed are mostly based on heuristic approaches. For example, the already mentioned centroid scheme [30] selects the neighbors based on the packet error rate. Other authors have proposed a scheme where the neighbors are determined by sorting the RSS values [99]. The following two sections use a parameter estimation approach and focus on the problem of computing an optimal threshold value when connectivity is derived from RSS measurements. The optimal threshold discussed is the value that minimizes the expected estimation error on the node positions.

4.1.2 Range-Free Localization as a Parameter Estimation Problem

The work described in this section aims at putting the choice of the connectivity model on a more rigorous footing and define a criterion of general applicability to convert the RSS values into proximity information. The starting point is the work of Patwari and Hero III [137], where localization is cast as a parameter estimation problem, and connectivity data is obtained by comparing the average RSS values against a threshold. After a preliminary introduction on the parameter estimation error, Section 4.1.4 introduces a simple example designed to explain the approach and obtain useful insights on connectivity based localization.

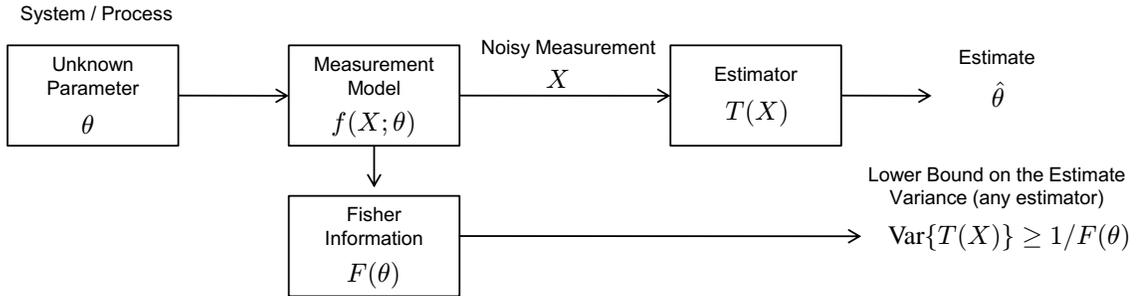


Figure 4.3: Schematic representation of the parameter estimation approach.

4.1.3 Preliminaries on the Parameter Estimation Approach

Many problems in science and engineering require estimation of parameters that describe some of the properties of a system or a process. Figure 4.3 shows a schematic representation of the case where the parameter θ has to be estimated using noisy measurements X . Depending on the initial information available, two different approaches can be used to define the *measurement model* that describes the dependence of X on θ . In the *Bayesian* estimation approach, θ is assumed to be a random value and the measurement model is defined by a conditional probability function $f(X|\theta)$. A priori information available on θ is expressed by a probability function $f(\theta)$. In the *Fisher* approach, which will be used in the following sections, θ is a deterministic but unknown parameter on which the probability function $f(X; \theta)$ depends.

Knowledge of the measurement model $f(X; \theta)$ can be used to design an estimator for θ , that is a function that uses X to produce an estimate of θ . The measurement models also serves to determine the uncertainty of the estimation process through the *Fisher information* defined as follows:

$$F(\theta) = E \left\{ \left[\frac{\partial}{\partial \theta} \log f(X; \theta) \right]^2 \right\}. \quad (4.1)$$

The function $F(\theta)$ is a measure of the amount of information that X carries about the unknown parameter θ [50]. If T is an estimator for θ , i.e. $\hat{\theta} = T(X)$, then the variance of T is bounded by the inverse of F :

$$\text{Var}\{T(X)\} \geq \frac{1}{F(\theta)}. \quad (4.2)$$

The inequality above, known as Cramér–Rao bound (CRB), sets a lower bound on the variance of *any* unbiased estimator³ that uses the measurement X . Notably, the CRB is not related to any particular estimation

³If $\hat{\theta}$ is an estimate of the unknown parameter θ obtained using T , $\hat{\theta} = T(X)$, then the estimator T is unbiased if $E\{\hat{\theta}\} = \theta$.

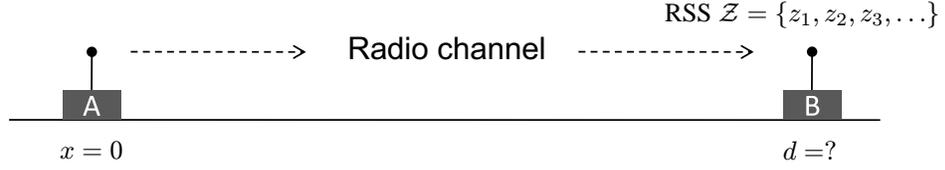


Figure 4.4: 1D localization: the distance of node 1 from the origin has to be estimated using connectivity information obtained by quantization of the RSS data.

technique, but it only depends on the measurement model.

In the following section, the parameter estimation approach and analysis of the Fisher information will be used to study a simple one dimension localization scenario with a single node in a one-dimensional space. The propagation model for the RSS signal is described in Section 4.1.5, while Sections 4.1.6 and 4.1.7 describe the quantization of the RSS values and the measurement model that relates connectivity measurement to the node position. Analysis of the Fisher information will then be used to investigate how to reduce the localization error by a proper choice of the quantization threshold.

4.1.4 Single Node Localization

Suppose two devices placed along a line as in Figure 4.4. Let the unknown parameter d denote the the position of node 1, which correspond to the the distance between the two nodes. The goal is to estimate d using connectivity information derived from RSS values. To enable localization, the two nodes exchange radio messages and collect a set (possibly empty) of RSS values: $\mathcal{Z} = \{z_1, z_2, z_3, \dots\}$. Let z be the average of the RSS value collected.

4.1.5 The log-normal shadowing model

This analysis assumes RSS measurements distributed according to the *log-normal shadowing model*, a propagation model that is widely used for link budget analysis in wireless communication. Adoption of this model is supported both by theoretical analysis of the RF propagation and by measurements in indoor and outdoor radio channels [143, 61, 23]. Another advantage of using the log-normal shadowing model is that it yields to analytically tractable results.

In condition of log-normal shadowing, the average RSS value z measured in dB (or dBm) is modeled as the outcome of a normal random variable Z with the following distribution:

$$Z \sim \mathcal{N}(P_r(d), \sigma_{\text{dB}}) \quad (4.3)$$

$$P_r(d) = P_0 + 10 n_p \log_{10} \left(\frac{d_0}{d} \right). \quad (4.4)$$

In the expression above, the term $P_r(d)$ denotes the expected value for the received power at a distance d when the received power between two nodes at the distance d_0 is P_0 . The parameter n_p is the *path loss exponent*, with typical values between 2 and 4 depending on the propagation environment. Finally, the standard deviation σ_{dB} models the variability measured between node pairs with the same separation distance, but at different locations (i.e. in different regions of the deployment area). Obstructions in the path between the nodes and reflections of the signal due to nearby obstacles can produce significant differences in the average received power measured by equidistant nodes. Typical values for σ_{dB} are between 3 and 12 dBm [143].

The average value z is considered instead of the values $\{z_1, z_2, z_3, \dots\}$ because a particular reading can be affected by large variability. Even if the nodes are static, movements of people, vehicles or other objects in the radio channel can cause RSS fluctuations that are uncorrelated with the node distance. Averaging the measured values reduces part of the signal variability and improves the localization results. On the downside, using the average RSS values forces the system designer to implement measurement protocols that exchange multiple messages and have a larger energy expenditure.

4.1.6 Threshold-based connectivity

According to the connectivity model discussed in Section 4.1.1, two nodes are connected if z is greater than a fixed threshold P_{th} and disconnected in the other case. The connectivity between two nodes is defined by a binary random variable C that takes the following values:

$$C = \begin{cases} 0 & \text{if } Z < P_{\text{th}} \quad (\text{nodes disconnected}), \\ 1 & \text{if } Z \geq P_{\text{th}} \quad (\text{nodes connected}). \end{cases} \quad (4.5)$$

According to (4.3) and (4.5), the probability of the event $C = 1$ (“nodes connected”) is the shadowed area in Figure 4.5a. The analytical expression for this probability is:

$$p = \Pr\{C = 1\} = 1 - G\left(\frac{P_{\text{th}} - P_r(d)}{\sigma_{\text{dB}}}\right), \quad (4.6)$$

where G is the CDF of a normal random variable $\mathcal{N}(0, 1)$. When the expected received power $P_r(d)$ equals P_{th} , the nodes are connected with probability $p = 0.5$. This condition occurs when the distance between the nodes equals the *threshold distance* d_{th} :

$$d_{\text{th}} = d_0 10^{\frac{P_0 - P_{\text{th}}}{10 n_p}}. \quad (4.7)$$

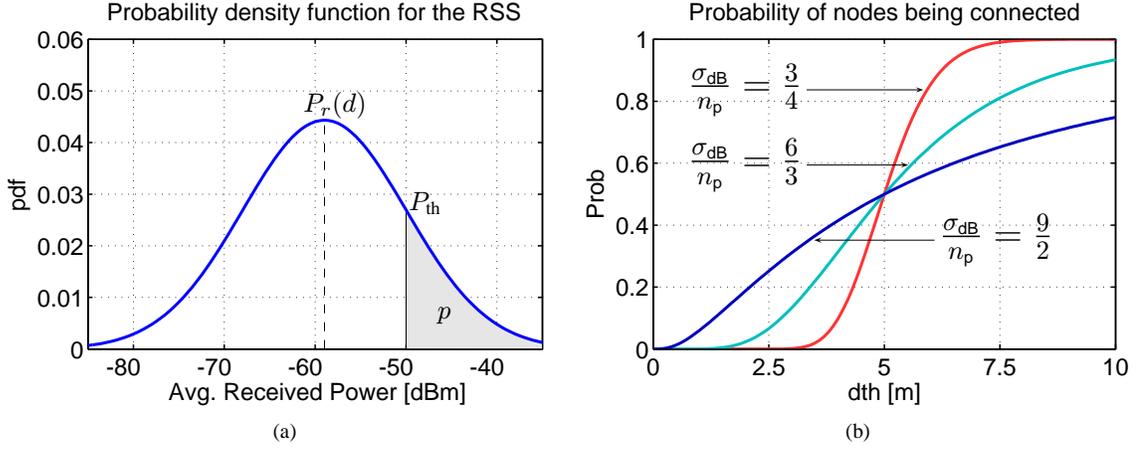


Figure 4.5: a) Probability density function of Z and probability p of detecting the node as “connected”; b) Probability of the event “nodes connected” as a function of d_{th} ($d = 5$ m).

Using the equation above, the probability of the event “nodes connected” can be expressed as a function of the node distance. Combining (4.4), (4.6) and (4.7), the probability p can be rewritten as [137]:

$$p = p(d, d_{th}) = 1 - G \left[K_c \log \left(\frac{d}{d_{th}} \right) \right], \quad (4.8)$$

where the constant

$$K_c = \frac{10}{\log 10} \frac{n_p}{\sigma_{dB}} \quad (4.9)$$

depends on propagation model’s parameters.

In the rest of this analysis, the problem of selecting the optimal threshold will focus on computing the value d_{th} . Fixed, the propagation model’s parameter, d_{th} can be converted into a RSS threshold value:

$$P_{th} = P_r(d_{th}) = P_0 + 10 n_p \log_{10} \left(\frac{d_0}{d_{th}} \right). \quad (4.10)$$

4.1.7 Fisher Information and Cramér-Rao Bound Analysis

The d_{th} value that minimizes the estimation error for the nodes’ distance is found by computing the Fisher information associated with the random variable C . The measurement model that described the dependence

of the connectivity data on the node distance is the *probability mass function* (pmf) of the random variable C :

$$f(c; d, d_{\text{th}}) = \begin{cases} 1 - p(d, d_{\text{th}}) & \text{if } c = 0, \\ p(d, d_{\text{th}}) & \text{if } c = 1, \\ 0 & \text{else.} \end{cases} \quad (4.11)$$

The pmf above satisfies the *regularity conditions* necessary to compute the Fisher information and ensure the CRB inequality [52]. In particular, the Fisher information is always defined (see below) and the *support* of $f(c; d, d_{\text{th}})$, i.e. the set of points where f is not zero, does not depend on d . The support of the function defined in (4.11) is the set $\{0, 1\}$.

The Fisher information for connectivity measurements (F_{con}) is a function of the parameters d and d_{th} defined as follows:

$$F_{\text{con}}(d, d_{\text{th}}) = E \left\{ \left[\frac{\partial}{\partial d} \log f(c; d, d_{\text{th}}) \right]^2 \right\} = \sum_{c \in \{0,1\}} \left(\frac{\frac{\partial}{\partial d} f(c; d, d_{\text{th}})}{f(c; d, d_{\text{th}})} \right)^2 f(c; d, d_{\text{th}}). \quad (4.12)$$

Since the inverse of the Fisher information defines the CRB, i.e. the lower bound on the estimation variance, the goal will be to find a d_{th} value that maximizes the amount of information available from connectivity measurements. To find the expression of F as a function of θ and d_{th} , (4.11) and (4.12) are combined. The resulting expression is:

$$F_{\text{con}}(d, d_{\text{th}}) = K_c^2 I_r(d, d_{\text{th}}) \left(\frac{1}{d} \right)^2, \quad (4.13)$$

where I_r is a term that depends on the ratio between θ and d_{th} :

$$I_r(d, d_{\text{th}}) = \frac{2}{\pi} \frac{\exp \left[-K_c^2 \log(d/d_{\text{th}})^2 \right]}{1 - \operatorname{erf} \left[\frac{K_c}{\sqrt{2}} \log(d/d_{\text{th}}) \right]^2}. \quad (4.14)$$

Figures 4.6a,b show the F_{con} computed as a function of the threshold distance d_{th} for different values of the ratio σ_{dB}/n_p and different node distances. As shown by (4.13), the information content of the measurements is inversely proportional to the square of the ratio σ_{dB}/n_p . Decreasing this ratio results in a sharper probability transition and larger values of the Fisher information (see Figure 4.5b and 4.6a). Intuitively, larger values of the parameter n_p imply a stronger correlation between the received power and the distance

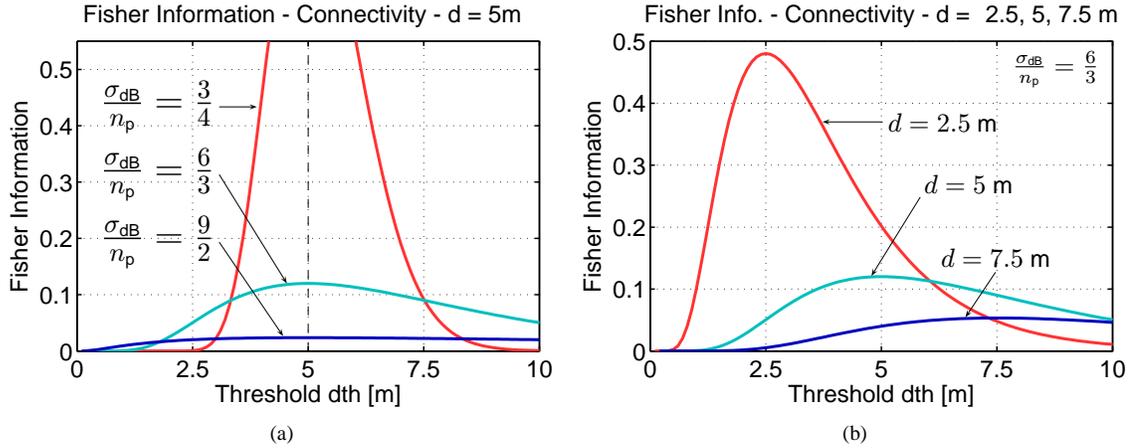


Figure 4.6: Fisher Information as a function of the threshold distance d_{th} : a) for nodes at $d = 5$ m and various value of the ratio σ_{dB}/n_p ; b) for nodes at distance $d = \{2.5, 5.0, 7.5\}$ m and $\sigma_{dB}/n_p = 6/3$.

between the nodes, which is a condition that causes the estimation error to decrease. In particular, the n_p term is a multiplicative factor in the expression $10 n_p \log_{10} (d_0/d)$ that measures the path loss between two nodes, where the path loss is the “signal” that carries information about the node distance. On the other hand, larger values of the parameter σ_{dB} pertain to environments where a strong shadowing noise increases the probability of measuring large deviations of the signal from the expected values. Since shadowing effects are not related to the node distance, their contribution should be regarded as a source of “noise” that reduces the accuracy of the estimation process.

4.1.8 Optimal Threshold for the 1D Case

While the parameters of the shadowing models depend on the radio environment and are out of a system designer’s control, the amount of information available can be maximized by properly choosing d_{th} . The plots in Figure 4.6a and 4.6b show that F always peaks when d_{th} equals d , and then it rapidly decreases to zero as the difference between d_{th} and d increases. *To reduce the estimation error, the threshold should be as close as possible to the true node distance (which is unknown).*

Threshold values with a large difference from d will reduce the amount of information available and result in less accurate estimates. For example, if the nodes are five meters apart and the chosen threshold is too low (e.g. $d_{th} = 2$ m), the two nodes will be disconnected with probability very close to one. The measurement carries little information about the true node distance because the nodes will almost always appear to be disconnected, no matter what the actual value of d is. From a localization point of view, we can

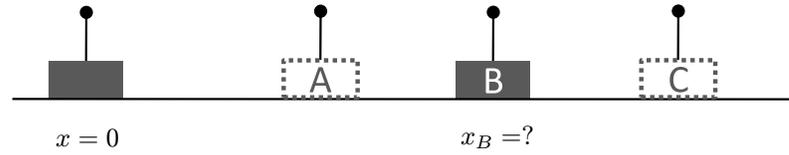


Figure 4.7: Localization example with a node placed in three possible positions.

only infer that the distance between the nodes is greater than 2 m ($d > 2$ m).

A similar situation occurs if the selected threshold is too large compared to the actual node distance (e.g. $d_{\text{th}} = 8$ m). The optimal choice is $d_{\text{th}} = 5$ m, which corresponds to nodes connected with probability $p = 0.5$. Also note that if the internode distance is increased, the optimal threshold is still achieved by setting $d_{\text{th}} = d$, but the information obtained from connectivity measurements decreases with the square of the distance between the two nodes (see Figure 4.6b). In other words, distance estimates for nearby nodes will be more accurate than distance estimates for nodes that are far from each other.

Example: Effect of Threshold Selection

An incorrect threshold selection will reduce the amount of Fisher information, increasing the error of a connectivity based scheme. To understand the effect of different thresholds, consider the following problem: Assume a node that can occupy three positions A, B, C at a distance 3, 5, and 7 m, respectively from a reference node (see Figure 4.7). The goal is to compute the true node position using connectivity measurements. Three quantization thresholds are available: low, medium and high with values P_L , P_M , and P_H . Which threshold will work better?

Assume B to be the true, but unknown node position. Figure 4.8 shows the distribution of the RSS values that would be measured for the node at different positions. Let p_a, p_b and p_c be the probability of measuring the node as connected. If the low threshold is used (i.e. $P_{\text{th}} = P_L$), it will be impossible to determine the position occupied by the node because all the three cases will produce connected measurements

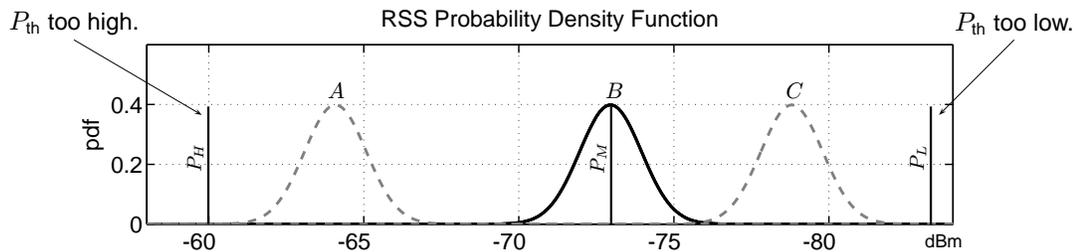


Figure 4.8: Distribution of the RSS values for a node at positions A, B, C, and three possible thresholds.

with probability close to one ($p_a \approx 1, p_b \approx 1, p_c \approx 1$). Similarly, using the high threshold will result in the same measurement (nodes disconnected) with probabilities close to one for all the three cases. Given the ambiguity in the measurements, the node could be placed either at A or C without changing the connectivity measurement. Neither P_L nor P_H allows a range-free scheme to determine the correct position.

The best threshold selection is $P_{th} = P_M$, which yields probabilities $p_a \approx 1, p_b = 0.5$, and $p_c \approx 0$. Since this threshold maximizes the probability to obtain different measurements for nodes at different positions, the expected localization error is lower than the error in the previous cases.

4.2 Optimal Threshold Selection in Collaborative Localization: The Optimal Connectivity (OC) Value

In general, a localization scheme is used to compute the positions of several nodes placed in 2D or 3D spaces. This scenario requires *collaborative localization* solutions (see Section 1.2). Even if a node is not in the radio range of any anchors, the proximity of other nodes (all placed at unknown locations) provide information to locate the node. This approach is also known as *multi-hop localization* because it supports localization of nodes placed several *hops* away from the anchors.

4.2.1 CRB Analysis

Consider a network with n nodes at unknown locations and m anchors. Similarly to the previous case, nodes collect RSS measurements and obtain connectivity values c_{ij} by comparing the average received power P_{ij} against a threshold P_{th} . Let \mathcal{C} be the set of all the random variables associated with the measurements:

$$\mathcal{C} = \{C_{ij} : C_{ij} \in \{0, 1\}, 1 \leq i, j \leq n + m\}. \quad (4.15)$$

The connectivity measurements are used to compute the n unknown node positions. The unknown coordinates can be arranged in a vector $\boldsymbol{\theta}$ with the following structure:

$$\boldsymbol{\theta} = \begin{cases} [\boldsymbol{\theta}_x, \boldsymbol{\theta}_y] & \text{if 2D localization,} \\ [\boldsymbol{\theta}_x, \boldsymbol{\theta}_y, \boldsymbol{\theta}_z] & \text{if 3D localization,} \end{cases} \quad (4.16)$$

where the vectors $\boldsymbol{\theta}_x, \boldsymbol{\theta}_y$ and $\boldsymbol{\theta}_z$ contain the unknown coordinates: $\boldsymbol{\theta}_x = [x_1, \dots, x_n], \boldsymbol{\theta}_y = [y_1, \dots, y_n]$ and

$\boldsymbol{\theta}_z = [z_1, \dots, z_n]$. Similar to the 1D case, analysis of the Fisher information and the CRB will be used to determine a threshold that minimizes the estimation error for $\boldsymbol{\theta}$.

In the case of collaborative localization, the measurement model is the *joint probability function*

$$f(\mathcal{C}; \boldsymbol{\theta}, d_{\text{th}}) = f(c_{11}, c_{12}, c_{13}, \dots; \boldsymbol{\theta}, d_{\text{th}}), \quad (4.17)$$

which relates the connectivity measurements to the node positions defined by $\boldsymbol{\theta}$ and the threshold distance d_{th} . Assuming independent RSS measurements, the joint probability (4.17) can be written as:

$$f(\mathcal{C}; \boldsymbol{\theta}, d_{\text{th}}) = \prod_{i,j=1}^{n+m} f(c_{ij}; \mathbf{v}_i, \mathbf{v}_j, d_{\text{th}}), \quad (4.18)$$

where \mathbf{v}_i and \mathbf{v}_j are the vectors with the coordinates of nodes i and j ; $\mathbf{v}_i = [x_i, y_i]^t$ or $\mathbf{v}_i = [x_i, y_i, z_i]^t$ depending on the dimensionality of the deployment space. Each pmf in (4.18) is similar to (4.11). In particular, two nodes i and j are connected with the following probability

$$p_{ij} = \Pr\{C_{ij} = 1\} = 1 - G \left[K_c \log \left(\frac{d_{ij}}{d_{\text{th}}} \right) \right], \quad (4.19)$$

where $d_{ij} = \sqrt{(\mathbf{v}_i - \mathbf{v}_j)^t (\mathbf{v}_i - \mathbf{v}_j)}$ is the Euclidean distance between the nodes. All the other symbols have the same meaning as in Section 4.1.

In the multi-parameter case, the information is measured by the *Fisher Information Matrix* (FIM) with the following elements

$$[F(\boldsymbol{\theta})]_{ij} = E \left\{ \frac{\partial}{\partial \theta_i} \log f(\mathcal{C}; \boldsymbol{\theta}, d_{\text{th}}) \frac{\partial}{\partial \theta_j} \log f(\mathcal{C}; \boldsymbol{\theta}, d_{\text{th}}) \right\}. \quad (4.20)$$

The FIM has $(2n \times 2n)$ elements for nodes placed in 2D spaces, and $(3n \times 3n)$ elements when localization computes 3D coordinates. Given the structure of the parameter vector defined in (4.16), the FIM is partitioned in sub-matrices $\mathbf{F}_{xx}, \mathbf{F}_{xy}, \dots, \mathbf{F}_{zz}$ with $n \times n$ elements each:

$$\mathbf{F} = \begin{cases} \begin{bmatrix} \mathbf{F}_{xx} & \mathbf{F}_{xy} \\ \mathbf{F}_{xy}^t & \mathbf{F}_{yy} \end{bmatrix} & \text{if 2D localization} \\ \begin{bmatrix} \mathbf{F}_{xx} & \mathbf{F}_{xy} & \mathbf{F}_{xz} \\ \mathbf{F}_{xy}^t & \mathbf{F}_{yy} & \mathbf{F}_{yz} \\ \mathbf{F}_{xz}^t & \mathbf{F}_{yz}^t & \mathbf{F}_{zz} \end{bmatrix} & \text{if 3D localization.} \end{cases} \quad (4.21)$$

More details on how to compute the FIM for the 2D case are given by Patwari and Hero III [137]. For the following analysis, it suffices to note that each sub matrix has elements similar to (4.13). For example, the elements of the sub-matrix \mathbf{F}_{xx} are:

$$[f_{xx}]_{ij} = \begin{cases} -K_c^2 \cdot I_r(d_{ij}, d_{th})(x_i - x_j)^2/d_{ij}^4 & (i \neq j) \\ K_c^2 \cdot \sum_{k=1}^{n+m} I_r(d_{ik}, d_{th})(x_i - x_k)^2/d_{ik}^4 & (i = j) \end{cases} \quad (4.22)$$

The sub-matrices \mathbf{F}_{yy} and \mathbf{F}_{xy} have a similar structure, but the terms $(x_i - x_j)^2$ are replaced by $(y_i - y_j)^2$ in \mathbf{F}_{yy} , and by $(x_i - x_j)(y_i - y_j)$ in \mathbf{F}_{xy} . Similarly, the terms in the sub-matrices \mathbf{F}_{xz} , \mathbf{F}_{yz} and \mathbf{F}_{zz} are: $(x_i - x_j)(z_i - z_j)$, $(y_i - y_j)(z_i - z_j)$ and $(z_i - z_j)^2$, respectively.

Anchor information contributes to the diagonal terms of each submatrix. At least three anchors are needed for localization in 2D, while four non-collinear anchor nodes are necessary for localization in 3D. Failure to include sufficient anchor information will cause the FIM to be rank deficient [117]. In this case, analysis of the CRB is possible using the Moore-Penrose pseudoinverse of the FIM [34]. The following analysis assume that the FIM is always invertible.

The inverse of the FIM bounds the covariance matrix of any unbiased estimator for $\boldsymbol{\theta}$ that uses observation from the set of random variables \mathcal{C} :

$$\text{Cov}\{T(\mathcal{C})\} \geq \frac{1}{\mathbf{F}}. \quad (4.23)$$

The diagonal elements of \mathbf{F}^{-1} are the lower bound for the variance on the node coordinates x_i, y_i and z_i : $\sigma_{x_i}^2 = [\mathbf{F}^{-1}]_{i,i}$, $\sigma_{y_i}^2 = [\mathbf{F}^{-1}]_{i+N,i+N}$, and $\sigma_{z_i}^2 = [\mathbf{F}^{-1}]_{i+2N,i+2N}$. The variance on the position of each sensor location is obtained by summation of the variance of the single coordinates:

$$\sigma_i^2 = \begin{cases} \sigma_{ix}^2 + \sigma_{iy}^2 & \text{if 2D localization,} \\ \sigma_{ix}^2 + \sigma_{iy}^2 + \sigma_{iz}^2 & \text{if 3D localization.} \end{cases} \quad (4.24)$$

If the same topology has to be localized in different environments (different realization of the random variables P_{ij} 's), then the terms (4.24) are a lower bound for the RMS error on the position of each node. Assuming that $\hat{\mathbf{v}}_i^{(1)}, \dots, \hat{\mathbf{v}}_i^{(k)}$ are K estimates for the position of node i , then:

$$\text{RMS}(i) = \sqrt{\frac{1}{K} \sum_{i=1}^K (\hat{\mathbf{v}}_i^{(i)} - \mathbf{v}_i)^t (\hat{\mathbf{v}}_i^{(i)} - \mathbf{v}_i)} \geq \sigma_i. \quad (4.25)$$

Notes on the Measurement Model

The model in this section assumes independent measurements between different pairs of nodes, which is a simplification of the reality. Shadowing of the RF signal is caused by static obstructions in the path between two nodes. Therefore, if two pairs of nodes share a large portion of the same physical path, it is reasonable to expect some correlation between the measurements taken over these two links. Correlation models describing shadowing across different links have been proposed in the literature (e.g. [63, 81, 62]), and, more recently, Patwari and Agrawal have studied the effects of correlated shadowing on the localization bounds [134]. The authors have shown that when shadow fading correlations are taken into account, the standard deviation bound measured by the CRB decreases of a few percentage points (between 2.4% and 4.5% in the example analyzed in their work). This evidence suggests that a localization scheme could improve its performance by taking into account the correlation existing between different measurements. However, since most of the available localization schemes are not designed to exploit this information, the assumption of independent measurements is reasonable to model the localization error achievable by range-free localization.

The measurement model in this section also assumes that *every* node makes measurements with *every other* node in the network. In practice, it may happen that two nodes are too far from each other to exchange messages and collect RSS information. In absence of external interferences, this situation occurs when the RF signal reaches the recipient with a power that is below the transceiver's sensitivity P_s .

The sensitivity P_s can be regarded as an implicit threshold set by the hardware. Since the probability of receiving messages with RSS lower than P_s is low, the threshold selection problem is meaningful only for values $P_{th} > P_s$. When nodes are unable to exchange radio messages, their RSS is lower than P_s and, consequently, lower than P_{th} . It follows that even pairs of nodes that are out of their radio range produce valid connectivity measurements. According to the discussion in this section and the model in Section 4.1.6, these nodes are always associated with the event "nodes disconnected".

4.2.2 The Optimal Connectivity (OC) Value

The lower limit on the variance of the node positions can be found by computing the inverse of the FIM. Since the Fisher information depends on the choice of the threshold distance d_{th} , the values (4.24) will also depend on d_{th} . The optimal threshold is chosen as the value that minimizes the average of the standard deviation of

the node positions:

$$d_{\text{th}}^* = \arg \min_{d_{\text{th}}} \text{CRB}_{\text{conn}}(d_{\text{th}}) \quad (4.26)$$

$$\text{CRB}_{\text{conn}}(d_{\text{th}}) = \frac{1}{n} \sum_{i=1}^n \sigma_i, \quad (4.27)$$

where the values σ_i depends on d_{th} as shown in the previous section. Note that the optimal values d_{th}^* can always be converted into a threshold P_{th}^* to be used for quantization of the RSS values.

From a theoretical point of view, the d_{th}^* value minimizes the variance of the estimation error. In practice, while it is not necessary that every range-free scheme will perform close to the CRB, it is reasonable to expect that the knowledge of d_{th}^* will be useful in avoiding the large localization errors caused by an improper threshold selection. To validate this assumption, two localization examples are considered. The DV-HOP [127], MDS [154], and the SOM schemes are used to localize two networks with nodes deployed in a 2D and a 3D space (see Figures 4.9a and 4.9c). The node positions were generated using the same noisy grid model described in Section 3.3.1.

Figures 4.9b and 4.9d show the localization errors averaged over 20 repetitions with different realization of the RSS values. The results support the choice of a quantization level based on (4.26): The three schemes achieve different localization errors, but in all of the cases, the minimum error is reached when the power threshold P_{th} is close to the P_{th}^* value that minimizes the CRB.

Optimal Connectivity

When a P_{th} value is used to quantize the RSS values, each node will be connected to a sub-set of its neighbors. Therefore, the effect of RSS quantization can be summarized by considering the resulting *network connectivity*, i.e. the average number of neighbors per node. In some cases, expressing the results as a function of the connectivity provides a more homogeneous comparison between different deployments. In fact, different from d_{th} and P_{th} , connectivity does not depend on the physical extension of the deployment area and the magnitude of the RSS values measured by the transceivers.

When the results are expressed as a function of the network connectivity, the optimal thresholds, d_{th}^* or P_{th}^* , correspond to an **Optimal Connectivity (OC)** value that minimizes the CRB (see Figure 4.10). From a system's designer point of view, knowing the OC value is important for two reasons:

1. At run-time, if the connectivity of the network to localize is too high, the localization error can be reduced by setting a RSS threshold that ensures an average connectivity equal to OC. According to the

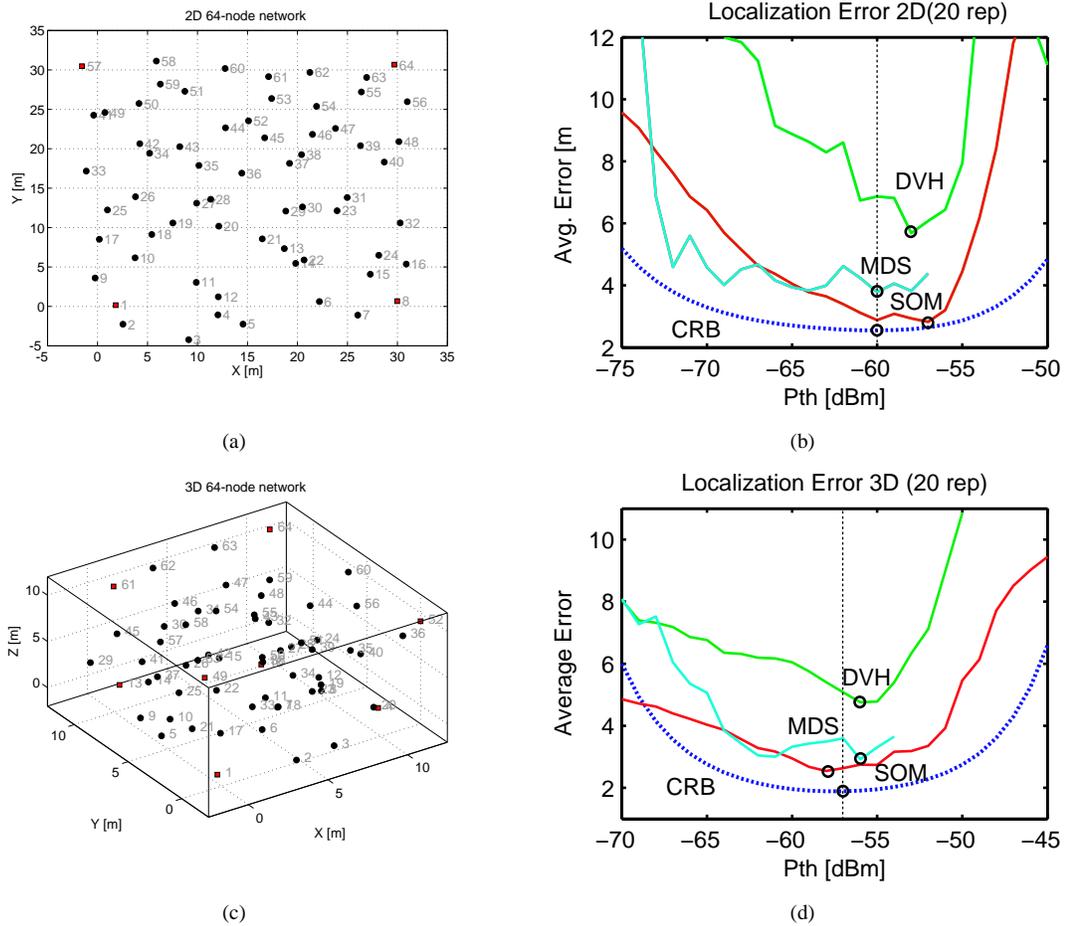


Figure 4.9: CRB and average localization error for localization in 2D and 3D spaces using three range-free schemes. **NOTE:** the CRB is compared to the average error only to illustrate the relation between its minimum and the error of the three schemes. Meaningful comparison between the absolute values should use the RMS localization error.

CRB analysis, this choice minimizes the localization error.

- At design time, the optimal connectivity can be used to guide the deployment of networks suitable for localization using range-free schemes. This design approach will be useful when using transceivers that do not support RSS readings. For example, the Bluetooth standard treats the RSS as an optional value whose purpose is only to define if the received power is within the *Golden Receive Power Range* [1]. Therefore, while not every Bluetooth implementation will accurately report RSS measurements suitable for quantization according to (4.5), the localization error can be still controlled by deploying networks with an average connectivity approximately equal to OC. A similar approach can be used when deploying networks that infer proximity constraints using RFID's.

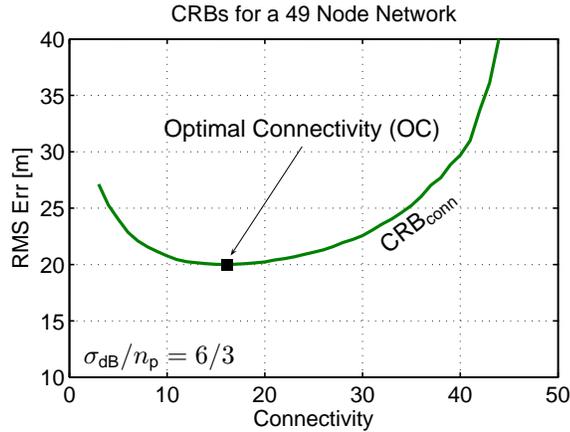


Figure 4.10: CRB for a 49 node network with four anchors on the corner of the deployment area. The Optimal Connectivity (■) value is marked on the plot.

Qualitative Analysis of the CRB for Range-Free localization

The existence of an optimal connectivity value can be explained intuitively by observing the convex shape of CRB_{conn} . Figure 4.10 shows that the CRB_{conn} increases when the network connectivity approaches values at the extremes of the range considered, i.e. the network connectivity is either very low or very high. Connectivity measurements are equivalent to knowledge of the neighbor set of each node. In the extreme case of a network with connectivity equal to zero, all the neighbor sets will be empty. In a fully connected network, all the neighbor sets will contain every node. In both cases, localization will not produce meaningful results because when all the nodes have the same neighbor sets, no information is available to discriminate their positions. Intuitively, between these two extreme values, there must be a connectivity value that minimizes the error.

The large error for extreme connectivity values can also be explained by analyzing the $F_{\text{con}}(d, d_{\text{th}})$ term derived for 1D localization. Only pairs of nodes with distance comparable to d_{th} contribute significant F_{con} values. When d_{th} is extremely small or extremely large, as in the cases discussed above, the total amount of information will be small because no pairs of nodes will have a distance similar to d_{th} .

For intermediate connectivity values, the choice of d_{th} determines which measurements are emphasized in the estimation process. Figure 4.11 shows the Fisher information available to estimate the position of a node in the center of the network. The nodes are plotted against a background that shows the $F_{\text{con}}(d, d_{\text{th}})$ at different distances and for two threshold values corresponding to $d_{\text{th}} = 15$ m and $d_{\text{th}} = 30$ m. Comparison of the plots indicates that *the choice of d_{th} determines a tradeoff between obtaining high-quality measurements from a few nearby nodes, or obtaining less valuable data for a larger number of nodes that are farther away.*

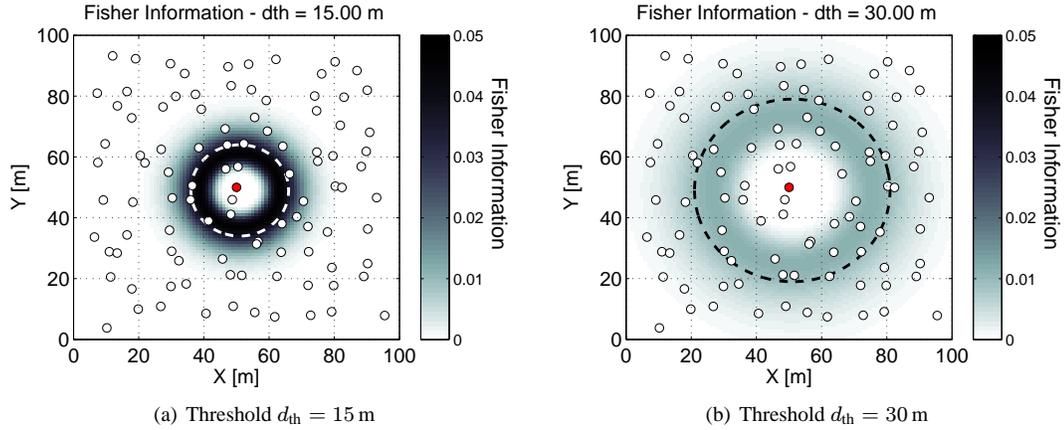


Figure 4.11: Effect of choosing two different threshold values. The background color indicates the Fisher Information $F_{con}(d, d_{th})$ at different distances from the node in the center. Darker colors correspond to a higher information content. Increasing d_{th} increases the number of nodes whose distance is similar to d_{th} , but since $F_{con} \propto 1/d^2$, these nodes contribute individually less information.

Therefore, the optimal connectivity is achieved by finding the d_{th} values that ensures the optimal trade-off between the information contribution of nodes at different distances.

Discussion

The qualitative analysis based on the results derived for the 1D case explains the existence of the OC value and the large errors for extreme connectivity values. While this analysis provides intuitive results, it should be noted that the CRB depends both on the distance and the geometrical configuration of the network nodes. A node having mostly collinear neighbors will have a large error even if the measurements have low noise. In cases similar to this one, performance degradation due to poor node placement is measured by the *Geometric Dilution of Precision* (GDOP) [48], and analysis of the error requires computing the CRB with all the network coordinates. However, when nodes have an (approximately) uniform distribution, the properties of the CRB can be understood using a qualitative analysis based on the results presented in Section 4.1.7. This is also the approach used to study some properties of the OC value in the following section.

4.2.3 Properties of the OC Value.

This section investigates how the OC values vary as the original network topology is transformed or the parameters of the propagation model change. The goal is to find an approximate OC value that can be computed without using the CRB. Besides the computational burden incurred in computing the inverse of a potentially large FIM matrix, the CRB analysis requires knowledge of the propagation model's parameters,

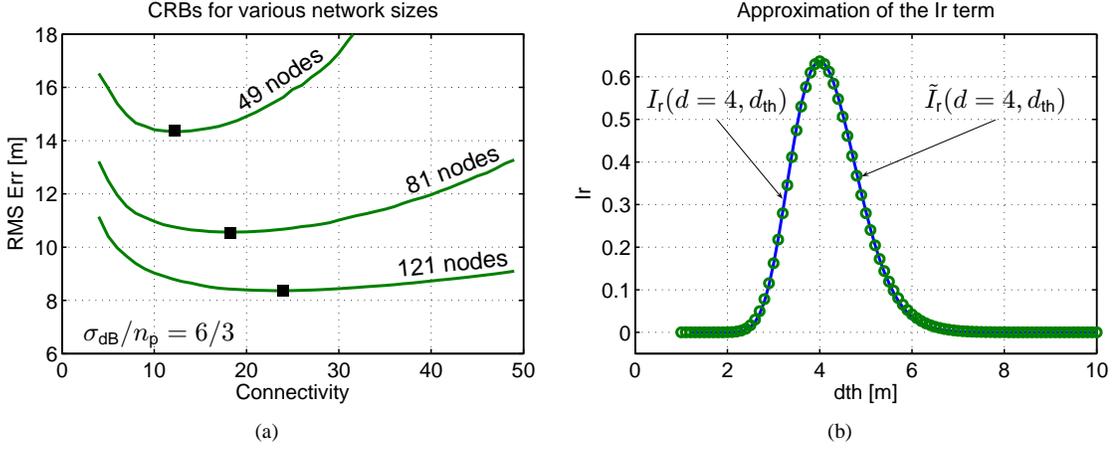


Figure 4.12: a) Optimal connectivity (■) computed for different network sizes; b) Approximation of the I_r term (solid line) using an exponential function (circles).

and, above all, the unknown node positions. As a result, a system designer trying to improve the performance of a connectivity-based scheme will not be able to compute the CRB to decide the optimal threshold.

In previous work, the CRB has been presented for localization using estimates of the inter-node distances. In that context, it was shown that the CRB is invariant under global translation, rotation or reflection of the network [34]. Except for the terms $I_r(\cdot, \cdot)$, the FIMs for distance and connectivity measurements have the same structure; therefore the same properties hold for connectivity-based localization. The next section analyzes the effect of various application parameters on the OC value. The parameters considered are: 1) the number of network nodes, 2) the ratio σ_{dB}/n_p , 3) the scaling factor for the node coordinates, and 4) the number of anchor nodes.

Number of Network Nodes

Figure 4.12a shows that OC increases with increasing values of the network size. To understand the effect of different number of nodes, consider the information available to estimate the position of a generic node. For simplicity, this analysis will consider a node in the center of the network similar to the case in Figure 4.11. Assuming independent RSS measurements, the available information is given by summation of the nodes' contribution at different distances:

$$F_{\text{tot}}(d_{\text{th}}) = \sum_i F_{\text{con}}^{(i)}(d_i, d_{\text{th}}) = \sum_i \frac{K_c^2}{d_i^2} I_r(d_i, d_{\text{th}}), \quad (4.28)$$

where d_i is the distance of the i^{th} neighbors. Again, this simplification does not account for geometrical configurations of nodes that could result in a large error. However, studying $F_{\text{tot}}(d_{\text{th}})$ does provide intuitive insights on the parameters that affect the OC value.

The optimal threshold d_{th}^* is the value that maximizes the available information. Since the analysis in previous sections has shown that the $F_{\text{con}}^{(i)}(d_i, d_{\text{th}})$ terms are non-negative, and that their value reduces to zero when $d \ll d_{\text{th}}$ or $d \gg d_{\text{th}}$, the optimal threshold is achieved in correspondence of a zero of the derivative:

$$\frac{\partial F_{\text{tot}}}{\partial d_{\text{th}}}(d_{\text{th}}^*) = \sum_i \frac{\partial F_{\text{con}}^{(i)}}{\partial d_{\text{th}}}(d_i, d_{\text{th}}^*) = 0. \quad (4.29)$$

Differentiation of the terms I_r 's in (4.28) yields a complicated expression, but the results can be simplified by considering an approximated form for I_r (see Figure 4.12b):

$$\tilde{I}_r(d, d_{\text{th}}) \approx \frac{2}{\pi} \exp\left(-\frac{\log(d/d_{\text{th}})^2}{K_a(\sigma_{\text{dB}}/n_p)^2}\right), \quad (4.30)$$

where $K_a \approx 0.13$ is a constant that was numerically determined using least square fitting. The intuition for using the above approximation is that the I_r term closely resembles a Gaussian kernel when its value are plotted on a logarithmic scale as a function of the ratio σ_{dB}/n_p . This property can be observed later in Figure 4.23a at page 83 in this Chapter. When the terms \tilde{I}_r are used in place of I_r , the terms in (4.29) have a more tractable expression:

$$\frac{\partial F_{\text{con}}^{(i)}}{\partial d_{\text{th}}}(d_i, d_{\text{th}}) = \frac{2K_c^2}{0.13(\sigma_{\text{dB}}/n_p)^2} \frac{1}{d_{\text{th}}d_i^2} \tilde{I}_r(d_i, d_{\text{th}}) \log\left(\frac{d_i}{d_{\text{th}}}\right). \quad (4.31)$$

Note that the sign of each derivative only depends on how each distance d_i compares against the threshold d_{th} :

$$\text{sign}\left(\frac{\partial F_{\text{con}}^{(i)}}{\partial d_{\text{th}}}(d_i, d_{\text{th}})\right) = \text{sign}\log\left(\frac{d_i}{d_{\text{th}}}\right) = \begin{cases} -1 & \text{if } d_i < d_{\text{th}} \\ 0 & \text{if } d_i = d_{\text{th}} \\ +1 & \text{if } d_i > d_{\text{th}}. \end{cases} \quad (4.32)$$

This result concords with the intuitive notion of the optimal threshold built so far. If all the neighbors are at distances less than d_{th} , the derivative of F_{tot} will be negative (see Figure 4.13a). The F_{tot} value can be increased by reducing d_{th} , i.e. moving it closer to the neighbors. If all the neighbors are at distances greater than the threshold, the derivative of F_{tot} will be positive. To obtain more information d_{th} needs to be increased.

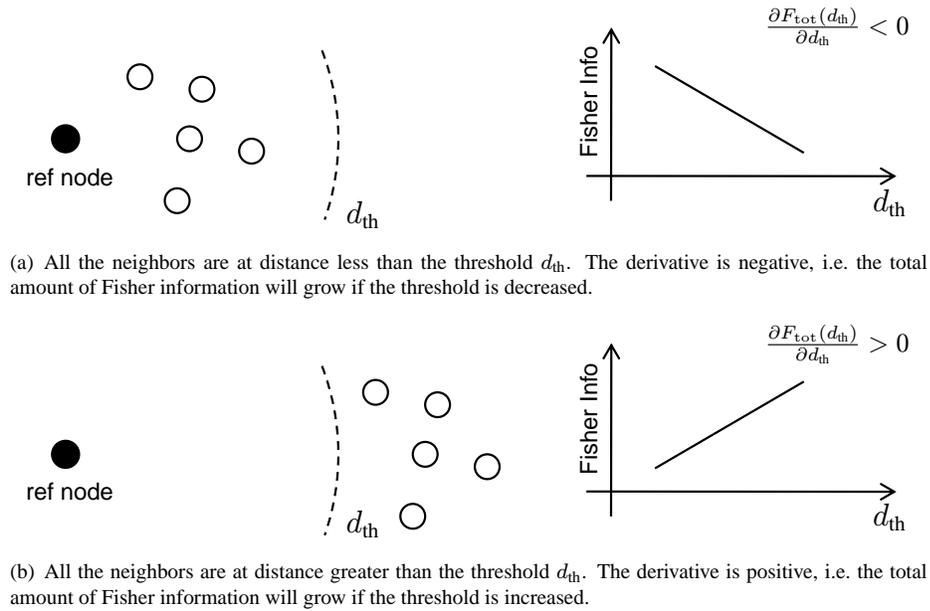


Figure 4.13: Derivative of the Fisher Information in two cases: a) all the neighbors of the reference node are at distances $d_i < d_{th}$: to increase the Fisher information d_{th} needs to be decreased; b) all the neighbors are at distances $d_i > d_{th}$: to increase the Fisher information d_{th} needs to be increased.

Consider now a network in which the threshold selected is optimal, i.e. $\partial F_{tot}/\partial d_{th} = 0$. If nodes are added at distance greater than d_{th} (i.e. without increasing the current connectivity level), the contribution of the new units will cause the derivative to become positive, thus violating the condition of optimality. To bring the derivative to zero, some nodes at distance less than d_{th} also need to be introduced, causing the optimal connectivity to increase. In conclusion, as shown in Figure 4.12a, the OC value will increase for increasing values of the network size.

Propagation Model Parameters σ_{dB}/n_p

Figure 4.14 shows the OC values computed for 64 node sample topology with different σ_{dB}/n_p values. There is no noticeable correlation between the noise in the RSS measurements and the position of the OC values. Again, analyzing of the derivative terms computed using the approximate function $\tilde{I}_r(d_i, d_{th})$ provides some insights on this property.

As discussed in the previous section, in condition of optimal connectivity, the negative contribution of the $\partial F_{con}^{(i)}/\partial d_{th}$ terms for $d_i < d_{th}$ must be balanced by the positive contribution of the terms with $d_i > d_{th}$. Consider a constant $p > 1$, and nodes placed at distance $d_1 = d_{th}/p$ and $d_2 = d_{th}p$. By replacing these

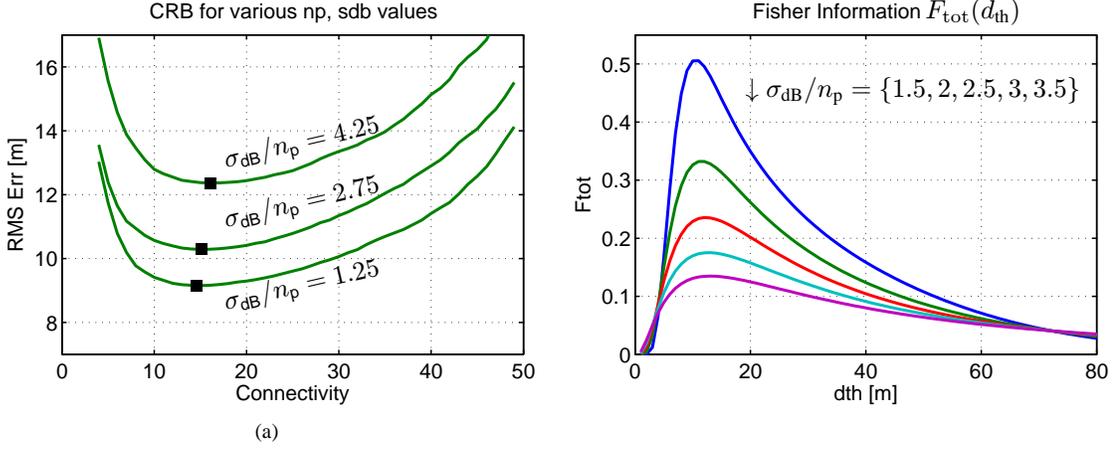


Figure 4.14: a) Optimal Connectivity (■) for a 64 node topology with different ratios σ_{dB}/n_p ; b) Value of the function $F_{tot}(d_{th})$ computed for a node in the center of a 64 node topology.

values in (4.31), it can be seen that the following equation holds:

$$\frac{\partial F_{con}^{(i)}}{\partial d_{th}} \left(\frac{d_{th}}{p}, d_{th} \right) = -p^4 \frac{\partial F_{con}^{(i)}}{\partial d_{th}} (p d_{th}, d_{th}). \quad (4.33)$$

For example, if $p = 2$ the presence of a node at distance $d_1 = d_{th}/2$ can be balanced by placing $2^4 = 16$ nodes at distance $d_2 = 2d_{th}$. The contribution $\partial F_{con}^{(i)}/\partial d_{th}$ of a node at distance $d_1 = d_{th}/3$ can be balanced by placing $3^4 = 81$ nodes at distance $d_2 = 3d_{th}$ and so on. Note that (4.33) holds for any value of the ratio σ_{dB}/n_p ; therefore if nodes were placed according the rule above, the OC would be exactly the same independently from σ_{dB}/n_p .

In typical WSN deployments, it is unlikely that the node distances will follow the distribution described. Depending on the value of the ratio d/d_{th} , altering σ_{dB}/n_p will cause some of the terms $\partial F_{con}^{(i)}/\partial d_{th}$ to grow more than others, possibly causing $\partial F_{tot}/\partial d_{th}$ to become different than zero. However, given the symmetry⁴ of the terms \tilde{I}_r around the value $d = d_{th}$, variations of σ_{dB}/n_p in the typical range measured in wireless applications seems not to alter significantly the position of the OC value. Figure 4.14b provides further support to this evidence by reporting the $F_{tot}(d_{th})$ values for a node in the center of a 64 node random deployment. Different values of the ratio σ_{dB}/n_p do not significantly alter the position of the F_{tot} maxima.

⁴Symmetry should be intended in the sense that $\tilde{I}_R(d_{th}/p, d_{th}) = \tilde{I}_r(d_{th} p, d_{th})$.

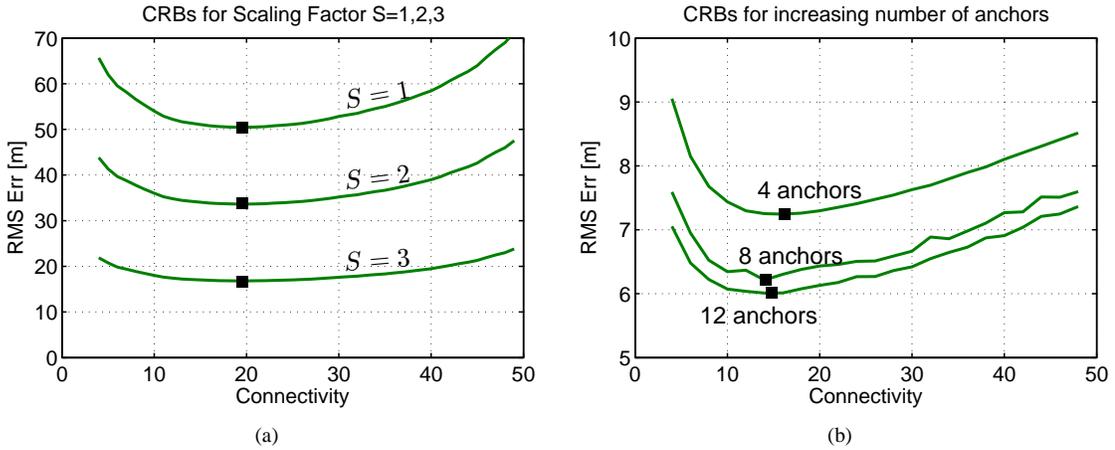


Figure 4.15: Optimal Connectivity: a) for different scaling factors; b) for increasing number of anchors.

Coordinate Scaling

The OC value does not change when the network coordinates are scaled by a factor S . This property follows from the equations that describes the Fisher information for connectivity measurements. Consider the term F_{con} discussed in Section 4.1.7, and assume that all the node distances are multiplied by a factor S . Also assume that d_{th} is scaled by the same factor, so the network connectivity remains constant. Under these conditions, the F_{con} term will be scaled by a factor S^{-2} :

$$F_{\text{con}}(Sd, Sd_{\text{th}}) = S^{-2}F_{\text{con}}(d, d_{\text{th}}). \quad (4.34)$$

When considering the multi-parameter case, scaling the network coordinates is equivalent to multiply the FIM matrix by a constant constant factor S^{-2} . As shown in Figure 4.15a, the position of the minima of the CRB_{conn} will not change.

Number of Anchor Nodes

Results of extensive simulations also show that increasing the number of anchor nodes cause the the CRB_{conn} to decrease, but without significantly affecting the OC position. As discussed in Section 4.2, anchor information contributes to the diagonal elements of the FIM. While a larger number of anchors will lower the error, there is no indication that this modification will alter the OC values. Figure 4.15b shows the CRB_{conn} for a sample topology with increasing number of anchors.

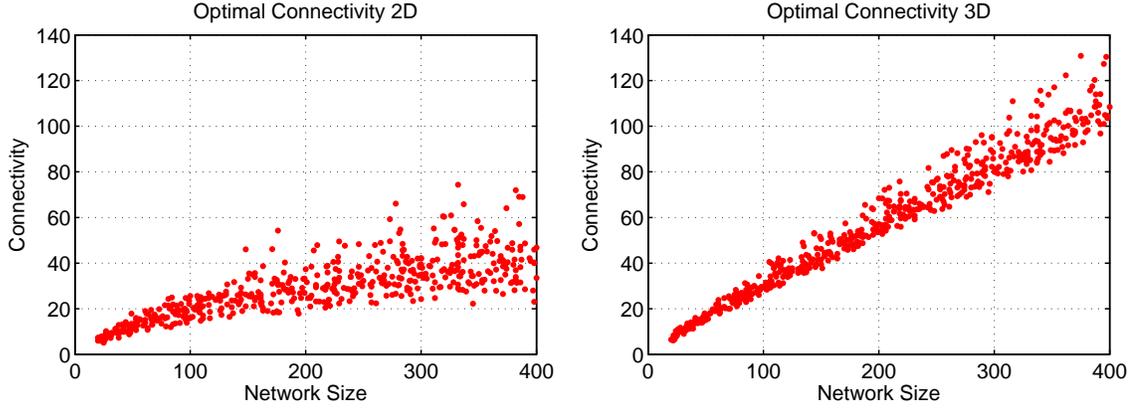


Figure 4.16: Simulation results: Optimal Connectivity for: a) 2D networks; b) 3D networks.

4.2.4 Approximation of the Optimal Connectivity Value

Previous sections have identified the network size as the only application parameter that sensibly affects the OC value. To model the dependence of OC on the number of nodes, the CRB_{conn} has been computed for a large number of simulated topologies with nodes placed in 2D and 3D spaces. Each case included about 500 random networks with a number of nodes between 20 and 400. The deployment areas were fixed: nodes were placed inside a square region $50 \text{ m} \times 50 \text{ m}$ for 2D networks, and in cube with side measuring 50 m for 3D networks. Four and eight nodes in the corner of the network were used as anchors for localization in 2D and 3D deployments respectively. For each network the parameters of the propagation model were uniformly sampled in the following intervals: $n_p \in [2, 4]$ and $\sigma_{\text{dB}} \in [3, 9]$ dBm.

Figures 4.16a,b show the simulation results. As expected, the OC value increases with increasing values of the network connectivity, but there is a noticeable difference in how the value grows in the two cases. An analysis of the derivative $\partial F_{\text{tot}} / \partial d_{\text{th}}$ similar to the one in Section 4.2.3 can help in understanding the differences between 2D and 3D networks. Again, the proposed analysis will consider a single node placed in the center of the network.

Assume a 2D network deployed in a circular region $A = \pi R^2$ with n nodes distributed according to a two-dimensional Poisson point process with density λ (see Figure 4.17). Assume also that c is the optimal connectivity value (i.e. $c = \text{OC}$) and d_{th} is the threshold value that realizes it. According to the hypothesis on the node distribution, the R and d_{th} and are related to the network size and connectivity by the following

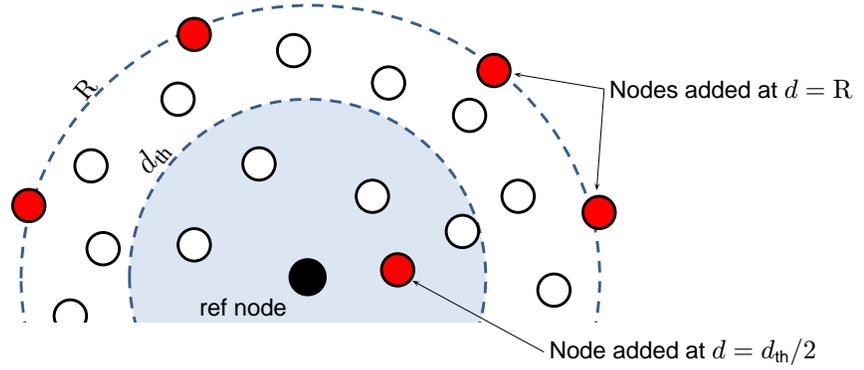


Figure 4.17: The network connectivity is increased by adding a node at distance $d_1 = d_{th}/2$. To maintain the optimality of the threshold d_{th} a number of nodes on the perimeter of the network need to be added.

equations:

$$(2D \text{ networks}) \quad R = \sqrt{\frac{n}{\lambda \pi}}, \quad (4.35)$$

$$d_{th} = \sqrt{\frac{c}{\lambda \pi}}. \quad (4.36)$$

To understand how the optimal connectivity and network size are related, consider increasing the connectivity by adding one device at distance $d_{th}/2$. According to the analysis in Section 4.2.3, the optimal connectivity is achieved when $\partial F_{tot}/\partial d_{th} = 0$; since adding a node at distance lower than d_{th} will cause the derivative to become negative, additional nodes at distance greater than d_{th} need to be added to maintain the optimality of d_{th} . For simplicity, it can be assumed that the new nodes are added at a distance $d = R$. Adding devices at distance $d < R$ would affect the network density and it would also complicate the analysis because some devices would fall inside the circle of radius d_{th} , thus altering the connectivity. The number of nodes to be added is found by evaluating the terms $\partial F_{con}^{(i)}/\partial d_{th}$ for nodes at distance $d_1 = d_{th}/2$ and $d_2 = R$. The contribution on the derivative for a node at distance d_1 is

$$DF_1 = \frac{\partial F_{con}^{(i)}}{\partial d_{th}}(d_1, d_{th}) = \frac{2K_c^2}{d_{th}} \frac{1}{d_1^2} \tilde{I}_r(d_1, d_{th}) \log\left(\frac{d_1}{d_{th}}\right). \quad (4.37)$$

Note that the expression above has been simplified by choosing a value of the ratio $\sigma_{dB}/n_p = 1/\sqrt{0.13} = 2.77$ dBm, so that the denominator inside the exponential in the \tilde{I}_r term is about one and can be omitted. The value for a node at distance $d_2 = R$ is equal to:

$$DF_2 = \frac{\partial F_{con}^{(i)}}{\partial d_{th}}(d_2, d_{th}) = \frac{2K_c^2}{d_{th}} \frac{1}{d_2^2} \tilde{I}_r(d_2, d_{th}) \log\left(\frac{d_2}{d_{th}}\right). \quad (4.38)$$

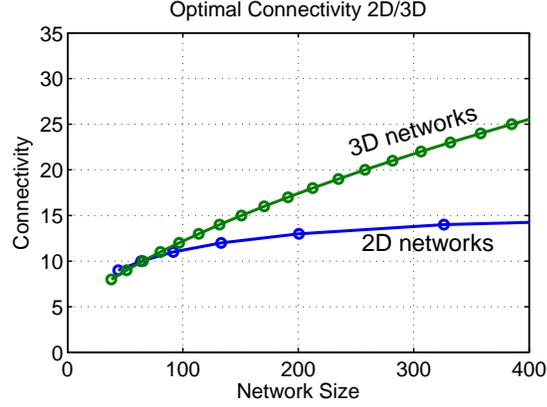


Figure 4.18: Trend of the theoretical optimal connectivity computed using an iterative approximation.

The number n_c of new nodes that need to be added at distance $d_2 = R$ must balance the contribution of the node at $d = d_{th}/2$:

$$DF_1 = -n_c DF_2 \quad (4.39)$$

Substituting (4.37) and (4.38) into (4.39) yields:

$$\begin{aligned} n_c = -\frac{DF_1}{DF_2} &= 4 \frac{n}{c} \frac{\exp(-\log^2 2)}{\exp(-\log^2(\sqrt{\frac{n}{c}}))} \frac{\log(2)}{\log(\sqrt{\frac{n}{c}})} \\ &= 8 \log(2) \frac{n}{c} \frac{\exp(\frac{1}{4} \log^2(\frac{n}{c}) - \log^2(2))}{\log(\frac{n}{c})}. \end{aligned} \quad (4.40)$$

The above expression computes the number of nodes n_c that need to be added to support an optimal connectivity $OC = c + 1$. It follows that $c + 1$ will be the optimal connectivity for a network with $n + n_c + 1$ nodes. Using $(c + 1)$ and $(n + n_c + 1)$ in place of c and n and repeating the same arguments, the network size values can be computed for increasing connectivity levels. This iterative scheme is described by the pair of equations:

$$c_{i+1} = c_i + 1 \quad (4.42)$$

$$n_{i+1} = n_i + 8 \log(2) \frac{n_i \exp(\frac{1}{4} \log^2(\frac{n_i}{c_i}) - \log^2(2))}{c_i \log(\frac{n_i}{c_i})} + 1 \quad (4.43)$$

The bottom plot in Figure 4.18 shows the pairs (n_i, c_i) with $i = 1, 2, 3, \dots$ computed using $c_0 = 9, n_0 = 44$ as a starting point (this values were determined by computing the CRB). The analysis in the 3D case is analogous, but the dependence of d_{th} and R on the current connectivity level c_i and network size n_i

is:

$$(3D \text{ networks}) \quad R = \sqrt[3]{\frac{3}{4} \frac{n_i}{\lambda \pi}}, \quad (4.44)$$

$$d_{th} = \sqrt[3]{\frac{3}{4} \frac{c_i}{\lambda \pi}}. \quad (4.45)$$

Taking into account the relations above, the number of nodes necessary to maintain an optimal threshold d_{th} is:

$$(3D \text{ networks}) \quad n_c = -\frac{DF_1}{DF_2} = 12 \log(2) \left(\frac{n}{c}\right)^{\frac{2}{3}} \frac{\exp\left(\frac{1}{9} \log^2\left(\frac{n}{c}\right) - \log^2(2)\right)}{\log\left(\frac{n}{c}\right)}. \quad (4.46)$$

The pairs (c_i, n_i) are also plotted in Figure 4.18 using $c_0 = 8, n_0 = 38$ as initial starting point. The two lines show a close similitude with the OC values experimentally computed in Figure 4.16. The main difference between the two cases is that in a 2D network, the minimum distance at which the nodes can be added is proportional to \sqrt{n} , while in the 3D case is proportional to $\sqrt[3]{n}$, where n is the number of nodes. It follows that a lower number of nodes is needed to support an increased connectivity in the 3D case. In other words, if the same number of nodes is added to the perimeter of a network, the optimal connectivity will grow faster for the 3D case.

Having explained the differences between the two cases, the OC value in Figure 4.16 can be approximated using simple functions. The OC values grows approximately as \sqrt{n} in the 2D case, while there is almost linear dependence between number of nodes and OC in the 3D case. The functions used for interpolation were empirically found and their coefficients were determined using least square fitting:

$$\tilde{OC}(N) = \begin{cases} -3.8290 + 2.3922\sqrt{n} & \text{if 2D localization} \\ 3.7055 + 0.2684 n & \text{if 3D localization} \end{cases} \quad (4.47)$$

Figure 4.19 shows the OC values together with the interpolation functions in the two cases.

4.2.5 Test Cases

Equation (4.47) implements a simple rule to determine what connectivity should be set when localizing a 2D or 3D network with a range-free scheme. Consider again the localization example discussed in Section 4.1.1. Equation (4.47) evaluated for a 2D network with 44 nodes indicates an optimal connectivity value equal to 12.04. For the network considered, this connectivity is achieved when $P_{th} = -54.22$ dBm. In Figure 4.20a, the optimal threshold (the vertical dashed line) is plotted together with the error of the DV-HOP, MDS and

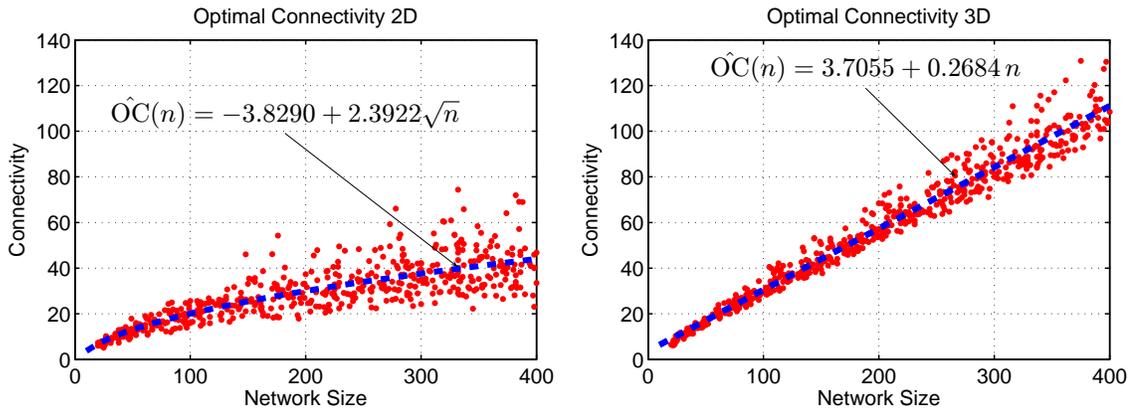


Figure 4.19: Using interpolation functions to approximate the OC values.

SOM algorithms. The plot also reports the CRB computed using the estimated values for the propagation model's parameters ($n_p = 1.7, \sigma_{dB} = 3.91$ dBm). The connectivity value given by (4.47) is close to the minimum of the CRB and close to the absolute minima of the MDS and SOM errors, thus validating the utility of the approximation found.

The second case study uses the RSS data from a 38 node network deployed in a 3D space [109]. The data is freely available on the ENALAB web site⁵. The optimal connectivity value found using (4.47) is 13.9, which for this network is achieved by setting a threshold $P_{th} = -45$ dBm. The error of the three localization algorithms for this network is reported in Figure 4.20b. Again, the estimated threshold results in an error that is close to the absolute minimum error for the three localization schemes.

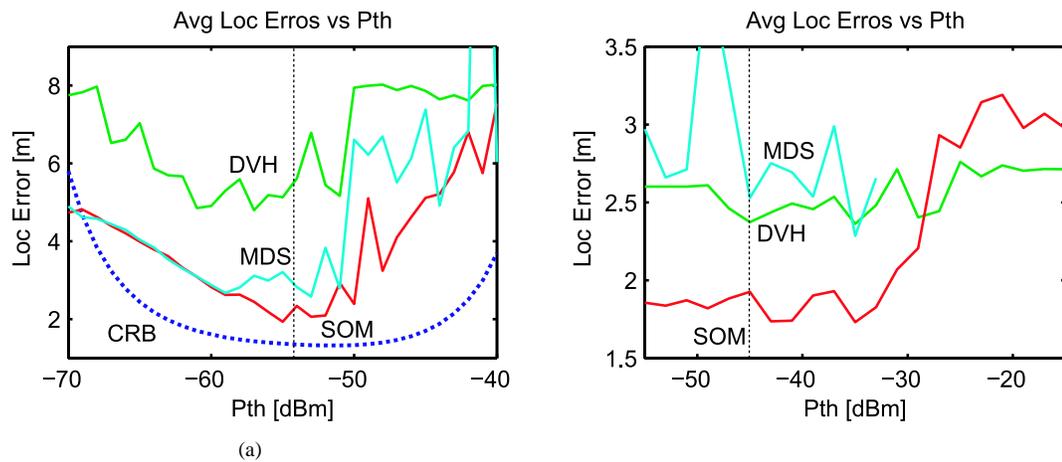


Figure 4.20: a) Localization error for the 44-node 2D network in [136]; b) localization error for the 38 node 3D network and in [109].

⁵http://www.eng.yale.edu/enalab/XYZ/data_set_1.htm

4.3 Comparison with the RSS-Ranging Approach

Previous sections have analyzed the proximity-based approach and the problem of selecting an optimal threshold value when connectivity data are obtained from RSS measurements. Alternatively, range-based localization can use the raw RSS values to estimate the inter-node distances (see Section 2.3.1). Similarly to localization based on radio connectivity, RSS ranging schemes are popular because no additional hardware is required on the nodes to be localized.

Since both the radio connectivity and the RSS-ranging approach are based on received signal strength values, a question arises. *Should the RSS data be used for range estimates, or should they be converted into connectivity information? Which approach works better?* As mentioned in the introduction, this problem has not been investigated in the literature.

In the next sections, the same parameter estimation approach of Sections 4.1 and 4.2 is used to compare the localization error of both range-free and range-based localization. The goal is to provide a practical rule to help system designers to identify the conditions under which an approach works better than the other. Similarly to the range-free case, the simple localization scenario presented in Section 4.1.4 will serve as a starting point for the analysis.

4.3.1 1D Node Range-Based Localization

Consider again the example in in Figure 4.4; in this case the distance between the two nodes has to be estimated using the unquantized RSS values. The measurement used in the estimation process is the value z computed by averaging the RSS values collected between the two nodes. As discussed in Section 4.1.5, the value z can be modeled as the outcome of a random variable Z with normal distribution (log-normal shadowing model). Under this assumption, an estimate of the distance can be computed using the *Maximum Likelihood Estimator* (MLE):

$$\hat{d}_{\text{ML}} = d_0 10^{(P_0 - z)/10n_p}. \quad (4.48)$$

If the path loss exponent n_p is known, the MLE provides a simple solution to convert RSS values into range estimates. Additionally, using (4.48) the estimation error can also be quantified. If the measurement is $z = P_r(d) + \delta$, where δ is a sample from the random variable $\Delta \sim \mathcal{N}(0, \sigma_{\text{dB}})$, then the error is:

$$e = \hat{d}_{\text{ML}} - d = d \left(10^{-\frac{\delta}{10n_p}} - 1 \right). \quad (4.49)$$

In absence of shadowing effects ($\delta = 0$), the MLE produces the correct estimates (i.e. $e = 0$). When $\delta \neq 0$, the error is proportional to the distance between the nodes; therefore, range estimates for nodes with a large separation distance are less accurate than range estimates for nodes that are close to each other.

Although the MLE for the node distance is readily available, the goal of this section is to compare the RSS ranging and the radio connectivity approaches on a more general basis. Again, analysis of the Fisher information will serve to derive results of general applicability and identify under which conditions the minimum expected error for one approach is lower than the other. As a result, the comparison will make it possible to select the localization technique capable of the lowest error.

In condition of log-normal shadowing model, the *measurement model* that relates the RSS value to nodes' distance is a normal distribution with pdf

$$f_Z(z; d) = \frac{1}{\sigma_{\text{dB}} \sqrt{2\pi}} \exp\left(-\frac{(z - P_r(d))^2}{2\sigma_{\text{dB}}^2}\right), \quad (4.50)$$

where $P_r(d)$ is the term defined by (4.4). The Fisher information is defined as

$$F_{\text{rss}}(d) = E \left\{ \left[\frac{\partial}{\partial d} \log f_Z(z; d) \right]^2 \right\}. \quad (4.51)$$

For the two nodes in Figure 4.4, substituting (4.50) into (4.51) yields

$$F_{\text{rss}}(d) = K_c^2 \frac{1}{d^2}, \quad (4.52)$$

where the constant K_c as the same value of (4.9). Figure 4.21 shows F_{rss} as a function of d for different values of n_p and σ_{dB} . The plots describe what was already seen in (4.49): The amount of information available to estimate d decreases for increasing values of the distance and increasing values of the ratio σ_{dB}/n_p .

Recall that the expression for the Fisher information computed for connectivity measurements is

$$F_{\text{con}}(d, d_{\text{th}}) = K_c^2 I_r(d, d_{\text{th}}) \frac{1}{d^2}.$$

This equation is similar to (4.52), but it contains the extra term $I_r(\cdot, \cdot)$ that depends on the ratio between the actual node distance and the threshold. Since the maximum value for the term I_r is achieved when $d_{\text{th}} = d$, F_{con} is approximately 37% lower than F_{rss} , even using the optimal threshold. In fact, $I_r(d, d_{\text{th}}) = 2/\pi \cong 0.63$ for $d_{\text{th}} = d$.

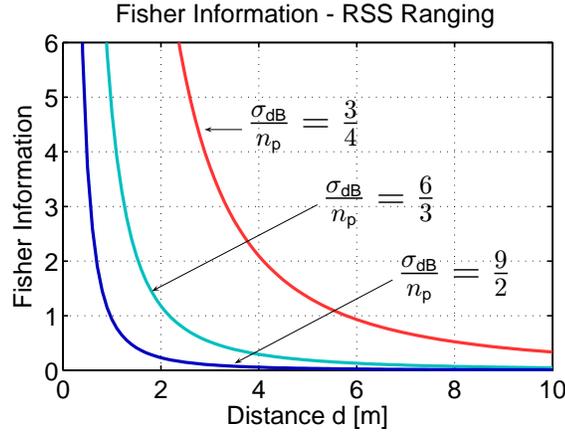


Figure 4.21: Fisher Information for RSS and connectivity measurements.

4.3.2 To Range Or Not To Range?

Comparison between F_{RSS} and F_{con} shows that RSS measurements always carry greater information content than connectivity ones; however, this is only true as long as the nodes are within the *radio range* of each other.

When nodes are within each other’s radio range, they can successfully exchange radio messages and z can be computed by averaging the values $\{z_1, z_2, z_3, \dots\}$. The value z can be used for range estimates using (4.48), or it can be used to derive connectivity information using (4.5). Depending on the choice of P_{th} , two nodes that are within each other’s radio range can be considered connected or disconnected.

On the other hand, when nodes are out of range the collection of RSS measurements is likely to fail. As discussed in Section 4.2.1 at pag. 64, when the power of the RF signal falls below the radio sensitivity, the demodulation of the incoming radio messages is likely to generate errors. Since the MAC layer of most transceivers is designed to silently drop packets containing errors, no RSS data will be made available to the application layer⁶. In this case, a range-based approach such as the MLE will not produce any position estimate (i.e. $F_{\text{RSS}}=0$). Instead, if a connectivity scheme is used, the occurrence of nodes that are out of range can be associated to the value “nodes disconnected”; therefore a position estimate is still possible ($F_{\text{con}}>0$).

The diverse nature of the measurements implies a fundamental difference between the two approaches. RSS ranging is more accurate when nodes are in the radio range of each other, but a connectivity scheme is naturally suited to localize nodes that are unable to communicate.

⁶Some transceivers (e.g. TI [5] and Jennic [2]) support test modes that allow the transmission of continuous waves. The power of these signals can be measured on the receiving unit without having to demodulate the signal; therefore RSS measurements are possible even below the radio sensitivity. A similar approach is used in the interferometric approach [112]. However, disabling the MAC layer in some units is likely to create interferences to nearby devices and can disrupt the functionality of the upper layer protocols (e.g. routing). Therefore the use of these special measurement modes is less appealing to collaborative schemes where multiple devices share a confined space.

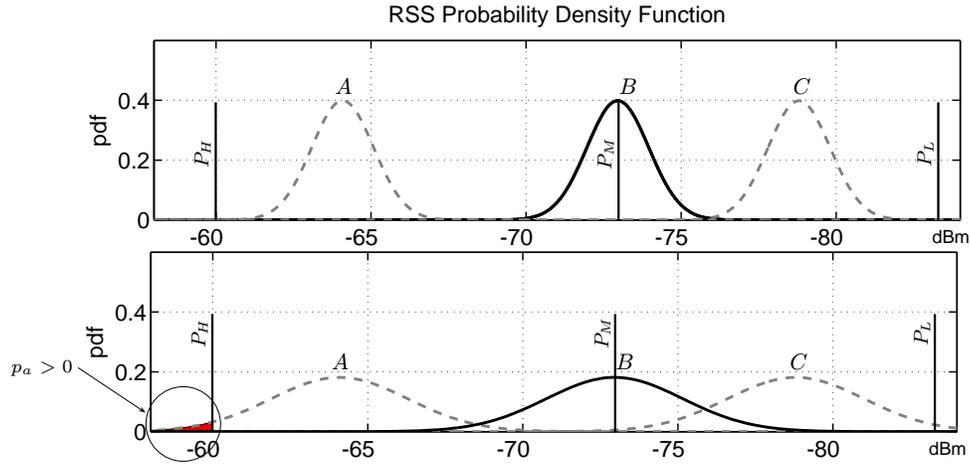


Figure 4.22: Top: Distribution of the RSS values for nodes at positions A, B, C, and three possible thresholds. Bottom: the same distributions computed with a larger ratio σ_{dB}/n_p . Note the the probability $p_a > 0$ that results from the noise in the measurements.

4.3.3 Effect of Shadowing on Range-Free and Range-Based Localization

High values of the ratio σ_{dB}/n_p degrade the quality of range estimates using RSS and have a similar effect on connectivity measurements (see Figures 4.21 and 4.6). However, while a strong shadowing variance has always detrimental effects on RSS range estimates, the occurrence of noisy measurements can sometime mitigate the effect of a wrong threshold selection. Consider again the localization example described in Section 4.1.8 (i.e. the example with the three nodes). Figure 4.22 compares the RSS distributions used in the previous example with more noisy distributions. Note that the threshold selections $P_{\text{th}} = \{P_L, P_H\}$, while still non-optimal, are not as ineffective as they were in the previous case. In both cases there is a nonzero probability to obtain a different measurement for at least one of the nodes at position A or C.

The effect of noise on connectivity-based localization can also be measured by evaluating F_{con} for different values of the ratio σ_{dB}/n_p . Figure 4.23a shows $I_r(d, d_{\text{th}})$ plotted as function of the ratio d/d_{th} for different values of the parameters n_p and σ_{dB} . Fixed the value d/d_{th} , I_r increases with increasing noise in the measurements. As a result, threshold selections that are ineffective for small values σ_{dB}/n_p will produce better results when σ_{dB}/n_p increases.

Figure 4.23b further illustrates the effect of noisy measurements on the threshold. The plots show the Fisher information values computed when nodes are five meters apart. The values $F_{\text{con}3}$, $F_{\text{con}5}$ and $F_{\text{con}7}$ measure the information when using a d_{th} equal to 3, 5 and 7 m respectively. The optimal threshold is $d_{\text{th}} = 5$ m; in fact, $F_{\text{con}5}$ is always greater than $F_{\text{con}3}$ and $F_{\text{con}5}$. However, when the ratio σ_{dB}/n_p is increased the differences between different choices become negligible.

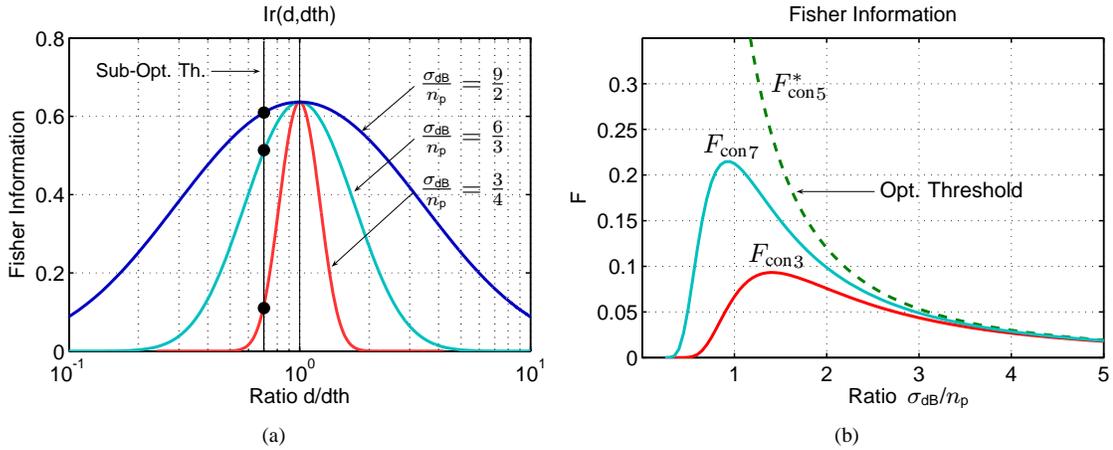


Figure 4.23: Effect of increasing the values of the ratio σ_{dB}/n_p the I_r term (left) and on the Fisher information values computed for various threshold selection (right).

4.3.4 Network Localization

In collaborative localization, the FIM and the CRB for a RSS ranging scheme are computed using the same approach described in Section 4.2. This section compares the localization limits for the range-free and the range-based case as a function of the network connectivity. The notations CRB_{RSS} and CRB_{conn} indicate the average value of the $2n$ coordinates' standard deviation for RSS ranging and localization based on radio connectivity respectively.

Figure 4.24 shows the CRBs for the same network used in Figure 4.10. This time the plot also reports the CRB_{RSS} . Different from the connectivity case, the CRB_{RSS} decreases monotonically with the connectivity. For RSS ranging localization, a given connectivity value, say ten, means that each node is in the radio range of other ten nodes; hence, ten range estimates are available to compute its position. As the connectivity increases, the number of measurements increases, causing the CRB_{RSS} to decrease.

4.4 Range-Free and Range-Based Collaborative Localization:

The Critical Connectivity (CC) Value

In addition to the OC value previously discussed, Figure 4.24 shows another important value: the *critical connectivity* (CC) value where the two CRB lines cross. For connectivity values below CC, CRB_{conn} is lower than CRB_{RSS} , implying that connectivity-based localization is potentially more accurate than RSS ranging, while the opposite is implied for values above CC. Assume a network with average node connectivity k .

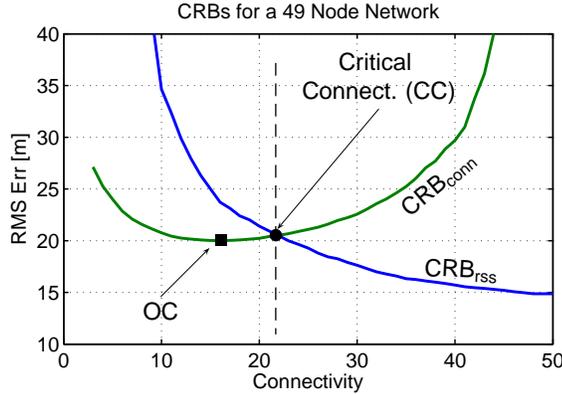


Figure 4.24: CRBs for a 49 node network with four anchors on the corner of the deployment area. The Critical Connectivity (●) and Optimal Connectivity (■) values are marked on the plot.

Comparing k against CC determines which localization approach should be used. A range-free scheme should be used if $k < CC$, and a range-based scheme if $k \geq CC$. According to the CRB analysis, this choice minimizes the expected localization error.

Similar to the OC analysis, the following sections study how relevant application parameters affect the CC value. The goal is to identify possible approaches to approximate this value without having to compute the two CRBs.

4.4.1 Properties of the Critical Connectivity

The CC value is studied for the same parameters previously used: 1) the number of network nodes, 2) the ratio σ_{dB}/n_p , 3) the scaling factor for the node coordinates, and 4) the number of anchor nodes. Since the CC is the intersection of the two CRBs, alteration in the relative position of the two limits will cause the CC value to change.

Number of Network Nodes

The CC value increases with the number of nodes in the network. Assume a network with average connectivity k . Assume now that the number of nodes is increased without changing k . If the deployment area is fixed, this is possible only if the communication range of each node and the RSS threshold are reduced. In this new network topology, a lower CRB_{rss} is expected because now the same number of measurements are available from nodes that are closer⁷. For connectivity-based localization, the CRB_{conn} decreases even more noticeably than the CRB_{rss} . In fact, for a range-free scheme the error reduces not only because nodes are

⁷ More in general, a proof by Patwari *et al.* [136] gives sufficient conditions for a decreasing CRB when new nodes are added to the network.

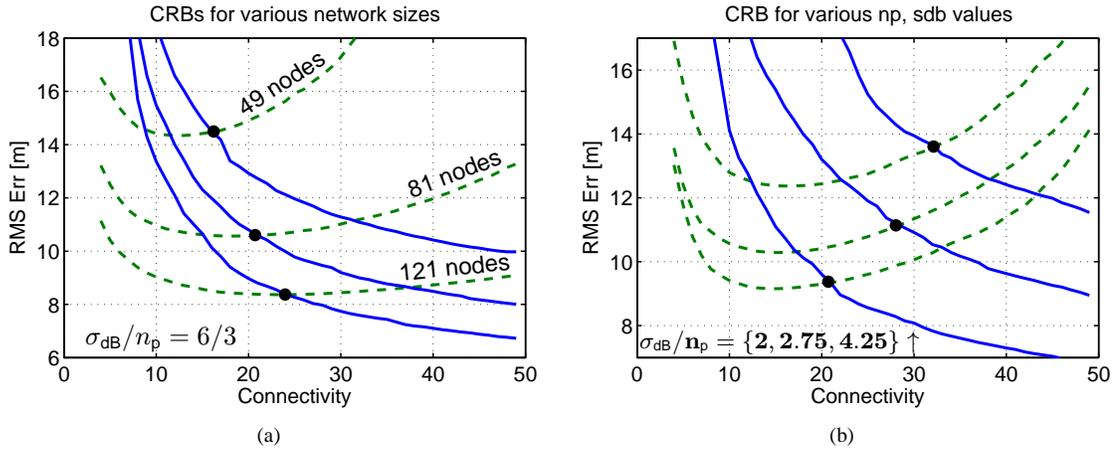


Figure 4.25: Critical Connectivity (●): a) for different network sizes; b) for increasing values of the ratio σ_{dB}/n_p .

closer to each other, but also because there are more measurements available from the disconnected nodes. Figure 4.25a shows the two CRBs computed for networks with 49, 81 and 121 nodes deployed in a square region with side equal to 100 m. While both CRBs decrease, the reduction is more evident for the CRB_{conn} . As a result, the CC value increases with the network size.

Propagation Model Parameters

The CC value also increases when the ratio σ_{dB}/n_p increases. As discussed in Section 4.1.7, this term describes the quality of the RSS the measurements. Increasing noise results in larger localization error; in fact, both the CRB_{rss} and the CRB_{conn} increases as shown in Figure 4.25b. But when using connectivity measurements, parts of the losses are compensated by the term $I_r(\cdot, \cdot)$, which was shown to increase with larger noise level (see Section 4.3.3). Since noise has a less severe impact on connectivity-based schemes, the CC increases with the ratio σ_{dB}/n_p (see Figure 4.25b).

Coordinate Scaling

Similar to the connectivity case, scaling the coordinates by a constant factor S has the same effect of multiplying the FIM by a factor S^{-2} . The expression for the Fisher Information for RSS ranging measurement is

$$F_{rss}(Sd) = S^{-2}F_{rss}(d). \quad (4.53)$$

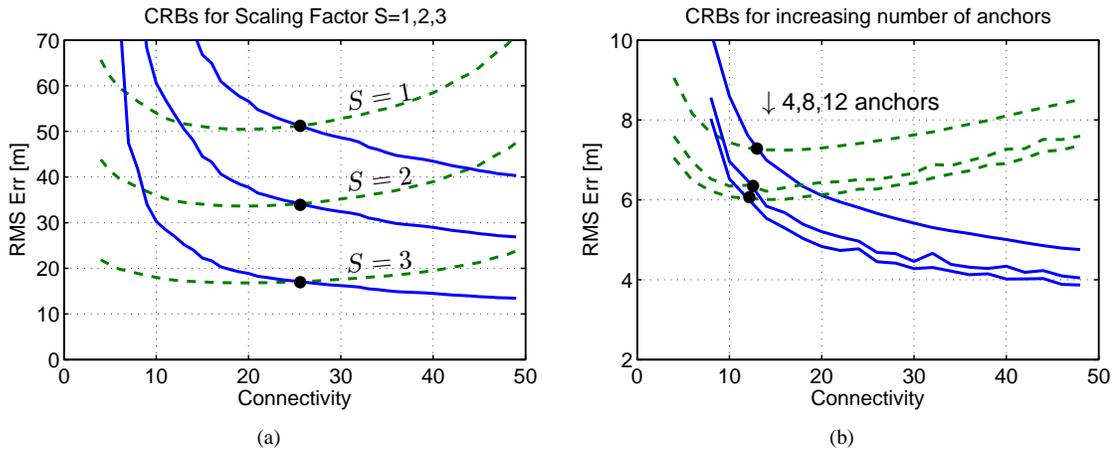


Figure 4.26: Critical Connectivity: a) for different values of the scaling factor: $S = \{1, 2, 3\}$; b) for increasing number of anchor nodes.

Since scaling the network coordinates has the same effect on the FIM elements for RSS and connectivity measurements, the relative position of the two CRBs will not change. This implies that the CC value also remains constant (see Figure 4.26a).

Number of Anchor Nodes

Results of simulations also shown that the CC value does not sensibly changes when a larger number of anchor nodes is used in the localization process. See Figure 4.26b.

4.4.2 Critical Connectivity Approximation

Table 4.1 summarizes the effect of the application parameters studied in previous sections on the OC and CC values. The CC value depends directly on the number of nodes in the network, and the ratio σ_{dB}/n_p . Again, a large number of simulated topologies is used to model the dependence of the CC value on these parameters.

Table 4.1: Effect of Parameters on CC and OC Values

Parameter	Effect on CC	Effect on OC
N. of Nodes	YES	YES
Ratio σ_{dB}/n_p	YES	Negligible
Scaling Factor S	NO	NO
N. of Anchors	Negligible	Negligible

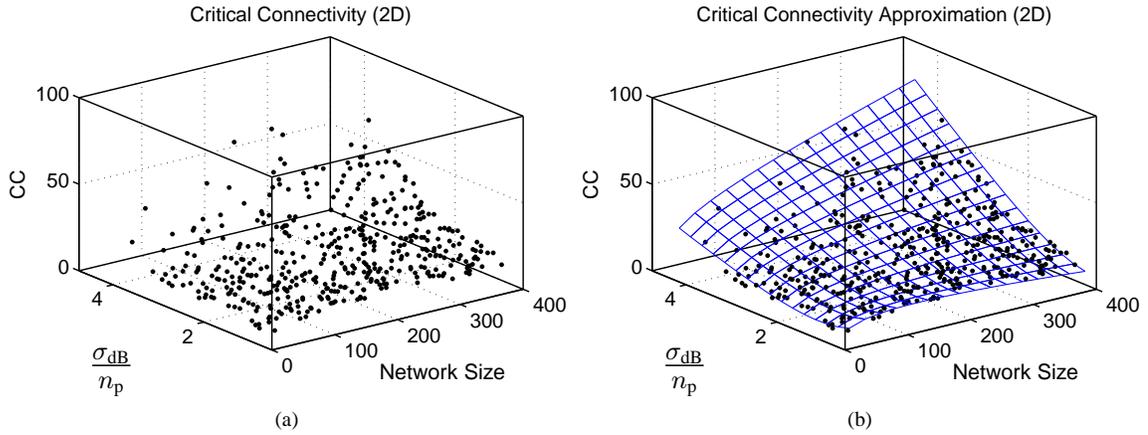


Figure 4.27: Simulation results for critical connectivity values and their approximation (2D networks).

The results are based on about 500 random generated using the same parameters described in Section 4.2.4. Figure 4.27a shows the simulation results. The CC values are plotted against the simulation parameters and appear to lie on a smooth surface. The values are interpolated using a function that is empirically found:

$$\begin{aligned} \tilde{C}C(n, r) = & a_0 + a_1 n + a_2 r + a_3 nr + \\ & + a_4 \log n + a_5 \exp(-r), \end{aligned} \quad (4.54)$$

where n is the number of nodes and r is the value of σ_{dB}/n_p . The values of the coefficient a_i , obtained by least squares fitting, are: $a_0 = -37.1022$, $a_1 = -0.0732$, $a_2 = 8.8506$, $a_3 = 0.0377$, $a_4 = 6.0667$, $a_5 = 41.8567$. The mean squared error between $\tilde{C}C(n, r)$ and the data points is equal to 6.18, while the average

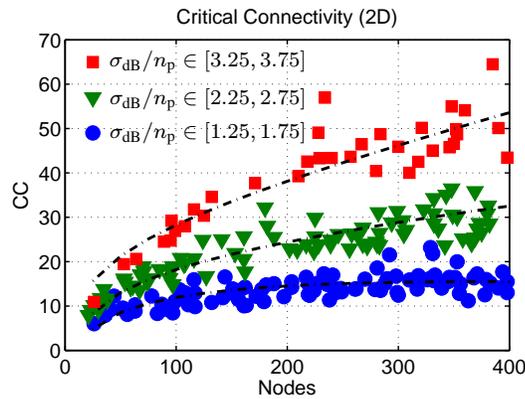


Figure 4.28: Approximation of the critical connectivity values and for different intervals of the values σ_{dB}/n_p .

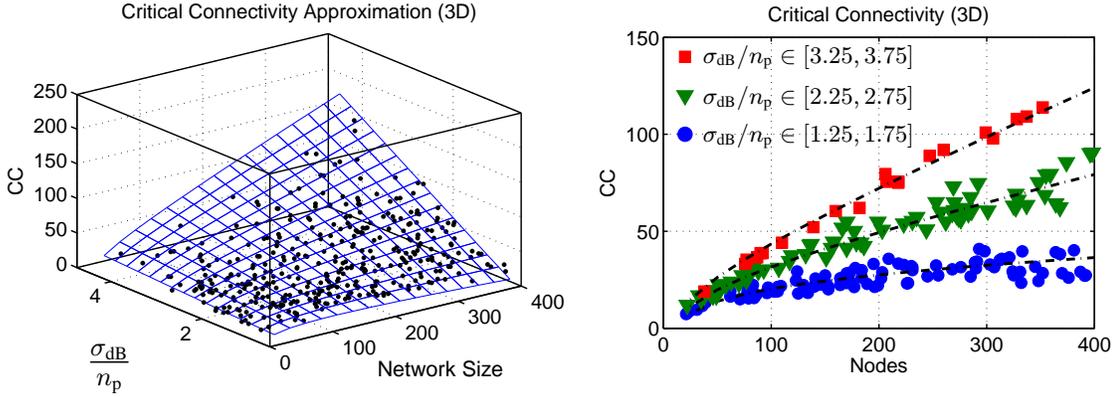


Figure 4.29: Simulation results for critical connectivity values and their approximation (3D networks).

error is equal to 1.88.

Figure 4.27b shows the interpolating surface (4.54) together with the data point. Figure 4.28 shows the CC values for different intervals of the ratio σ_{dB}/n_p . The dotted lines are computed using (4.54) for r equal to the central value of the σ_{dB}/n_p ranges considered. For low values of σ_{dB}/n_p , the CC value stabilizes around 15. As the ratio σ_{dB}/n_p increases, however, there is an higher correlation between the network size and CC values; therefore range-based schemes are beneficial only in highly connected networks. These results confirm the observations of other authors, who have occasionally noted that connectivity-based schemes outperform range-based ones in conditions of low connectivity [35] or when the ranges are estimated using noisy measurements [20, 129].

Results of simulations with network deployed in 3D spaces are similar to the 2D case. Figure 4.29 shows the point and the approximation surface used to interpolate them. The values of the coefficient a_i , obtained by least squares fitting, are: $a_0 = -20.5671$, $a_1 = -0.1480$, $a_2 = 1.0249$, $a_3 = 0.1092$, $a_4 = 8.0842$, and $a_5 = 2.0966$.

4.4.3 Test Case

Consider the 100 node network of Figure 4.30 with parameters $n_p = 3$ and $\sigma_{dB} = 8$ dBm. Application of (4.54) yields : $\tilde{CC}(100, 8/3) = 20.07$ (the exact value found using the two CRBs is 22). According to the proposed analysis, a connectivity based scheme should be used for connectivity values below 20.07, and a range-based scheme when the network's connectivity is above 20.07.

To validate the choice suggested by the approximated CC value, the node positions are computed

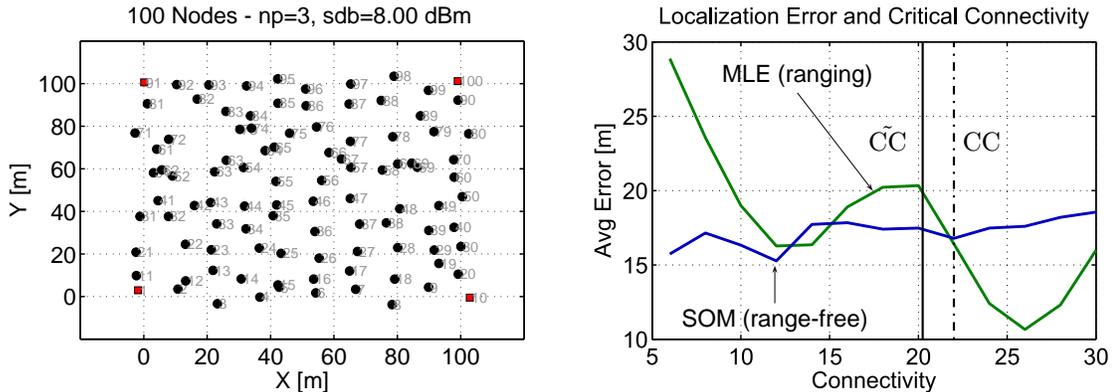


Figure 4.30: 100 node network test case.

using two algorithms. The first one is the SOM-A localization scheme, which has shown to perform well for low connectivity values. The other one is a range-based scheme that computes the MLE using gradient descent⁸ [136]. As shown in Figure 4.30, *a posteriori* analysis of the error confirms the choice made by using (4.54). For connectivity values lower than 20, the range-free scheme’s error is lower than the MLE’s error; the opposite is true for connectivity above 20.

4.5 Related Work

Over the past few years, analysis of the CRB have been used by a number of authors to characterize the error bound of localization algorithms, especially when using range measurements (angle or distances) affected by Gaussian noise. Moses et al. [117] have derived the CRB for localization based on signals emitted by a set of sources, and nodes can measure the Time of Arrival (ToA) or the Angle of Arrival (AoA). A study of the CRB under various conditions of node and beacon density has been proposed by Savvides et al. [8]. Wang et al. have defined a Bayesian Bound (BB) that is the covariance of a posterior distribution computed from the sensor observations [168]. This bound is equivalent to the CRB for measurements with Gaussian error, but it is computationally less demanding. Analysis of the CRB has been proposed by Patwari et al. for collaborative localization using distance estimates obtained by ToA and RSS [136], and for localization using angle estimates [135]. Localization using connectivity information or quantized RSS levels has been studied by Patwari and Hero III [137]. The idea to obtain connectivity data from RSS value has been also used by Li et al. to implement a *Partial Range Information* (PRI) scheme that derives “sub-hop” information useful in improving the localization accuracy [99]. This idea is somewhat similar to the one proposed in Section 4.2,

⁸We use the output of the range-free scheme as initial position for the gradient descent.

since the OC approach also tries to improve the localization accuracy by choosing a threshold for the RSS values.

Chapter 5

Localization in Heterogeneous Scenarios

Using SOM

Implementing a localization service for ad hoc networks is a challenging task. Sometimes the nodes are deployed in sparse topologies, while other times they are densely packed inside a building. Some environments are relatively uncluttered, while others have obstacles that impede the node placement and strongly affect the radio signal. To address the problem of localization in heterogeneous scenarios, the SOM schemes are validated using new extensive simulation sets based on log-normal shadowing model.

The simulations in this chapter show that the SOM techniques is be robust to conditions of strong shadowing of the RF signal and produces accurate results in a variety of simulated environment with nodes places in 1D, 2D and 3D configurations. However, the SOM's error also confirms the theoretical results discussed in Chapter 4: Range-free localization is not effective in densely-deployed networks. To avoid this limitation, insights gained from the theoretical analysis in Section 4.3 and 4.4 are used to devise an improved version of the algorithm (SOM-R) capable of combining connectivity measurements with RSS values. The new scheme is validated using RSS traces collected from wireless devices in three different environments. Results show a localization error that is substantially lower than other SOM variants and practically independent from the connectivity of the network being localized. Similar results are achieved in networks with anisotropic layouts, which are typically harder to localize.

5.1 Performance of the SOM based Localization under Log-Normal Shadowing and Comparison with the CRB

The *ideal radio model* used in Chapter 3 provides an intuitive abstraction useful in simulation studies, but it does not adequately capture the random nature of wireless communication. Multi-path fading due to reflection, diffraction and scattering of the RF signal causes variations in the received power and ultimately affects the capacity of the recipient to correctly decode a radio message.

This section evaluates the performance of SOM-A, MDS and DV-HOP using a new simulation set based on the log-normal shadowing model. The data used in the simulation are generated according to a two-step process: 1) For each pair of nodes, the average RSS values are sampled according to the log-normal distribution described by (4.3) and (4.4). 2) Connectivity data are obtained by binary quantization of the RSS values (see Section 4.1.6). To allow a meaningful comparison with the CRB, the simulation is repeated fifty times, using different realizations of the random variables that model the RSS values. For each trial, the localization results are evaluated at different connectivity levels obtained by adjusting the quantization threshold P_{th} in (4.5).

The performance metric used to evaluate SOM-A, MDS and DV-HOP is the average *Root Mean Square* (RMS) error defined as:

$$\text{RMS Err} = \frac{1}{n} \sum_{i=1}^n \sqrt{\frac{1}{K} \sum_{k=1}^K \left(x_i - \hat{x}_i^{(k)} \right)^2 + \left(y_i - \hat{y}_i^{(k)} \right)^2}, \quad (5.1)$$

where K is the total number of repetitions (fifty in this case), n is the number of nodes (excluded the anchors), (x_i, y_i) are the true node coordinates, and $(\hat{x}_i^{(k)}, \hat{y}_i^{(k)})$ are the coordinates computed using one of the localization algorithm in the k^{th} trial. The RMS Error is compared against the average value of the node positions' standard deviation computed using the CRB.

Additionally, each simulation is repeated in different shadowing conditions, with different values of the ratio $\sigma_{\text{dB}}/n_{\text{p}}$:

- 1) **Low noise:** $\sigma_{\text{dB}}/n_{\text{p}} = 3/4$ dBm.
- 2) **Medium noise:** $\sigma_{\text{dB}}/n_{\text{p}} = 6/3$ dBm.
- 3) **High noise:** $\sigma_{\text{dB}}/n_{\text{p}} = 9/2$ dBm.

As discussed in the previous chapter, the ratio $\sigma_{\text{dB}}/n_{\text{p}}$ describes the quality of the RSS measurements and

for MDS that performs worse than the other two solutions for high noise values (see Figures 5.1d). The error generated by SOM, which is inherently a stochastic scheme, and therefore less sensitive to measurements errors, degrades more gracefully as the noise increases.

For large connectivity values, the error plots are lower than the CRB. Increasing the threshold value excessively causes the FIM information to become rank deficient, and some of the components of the CRB go to infinity. In practice, the localization error of the three scheme is always finite, therefore it might be lower than the CRB when then network connectivity approaches the network size.

5.1.1 Localization in 1D and 3D spaces

In addition to the canonical application of localization in two dimensions, the simulations also evaluate the three schemes when nodes are placed in 1D and 3D spaces. Ad-hoc networks with linear configurations of nodes find application in traffic monitoring along highways and perimeter control. 3D deployments are found in asset tracking applications for large warehouse, or when instrumenting multistory buildings for ubiquitous computing.

All the three localization algorithms can be easily modified to work in dimensions different from two. In particular, the only modifications required by SOM are the use of weights with a different dimensionality and a corresponding change in the sampling space. Points are sampled from a line for localization in 1D, and from a cube for localization in 3D. In both cases, the training points for the SOM are sampled from uniform distributions with extension computed using (3.8).

Figure 5.2 shows the test topologies used and the results for medium noise ($\sigma_{\text{dB}}/n_p = 6/3$ dBm). The results for other noise levels are qualitatively similar to the 2D case. The only exception is the MDS scheme, which produces a large localization error in all the 1D configurations tested. In contrast, both SOM and DV-HOP achieve an error close to the bound for low connectivity values.

5.2 Localization in Dense Networks

Both simulation results and CRB analysis show a large localization error when the connectivity reaches values close to the network size (see Figures 5.1 and 5.2). When most of the nodes are in the radio range of each other, connectivity data are of scarce utility in determining the node positions. In particular, the extreme case of a fully connected network is the result of an improper threshold selection, which results in a small amount of Fisher information and large estimation errors (see Section 4.2.2).

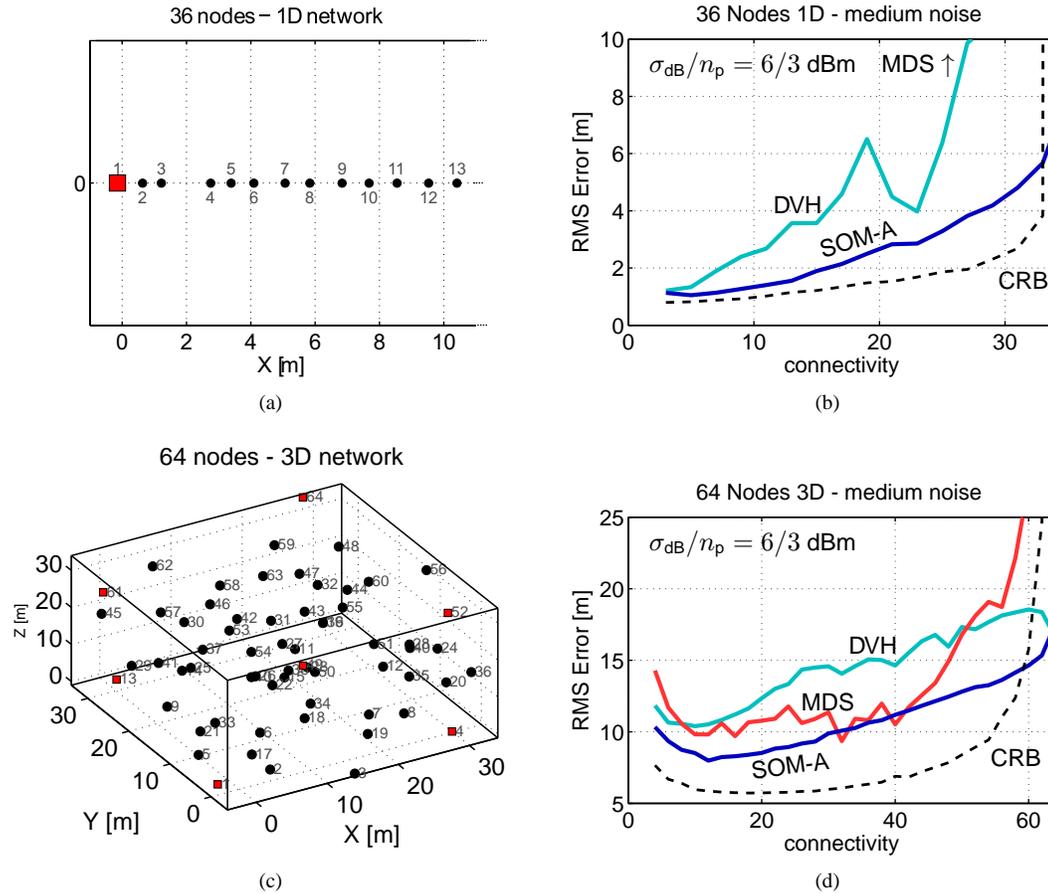


Figure 5.2: a,c) 1D and 3D sample topologies (red squares are anchor nodes); b,d) average RMS error achieved by SOM-A, MDS, DV-HOP for different connectivity levels and comparison with the CRB value.

When connectivity data is obtained from RSS values, any range-free scheme can be used to localize dense networks by applying the optimal threshold value discussed in Section 4.2. This approach will also work with the SOM localization algorithm. In the test cases analyzed, it was shown that the error achieved in correspondence of the optimal threshold is close to the absolute minimum error achieved by SOM (see Figures 4.20a and 4.20b at page 78).

The analysis in the previous chapter also shows that range-free algorithms are better suited to localize sparse networks, while range-based approaches work better in dense deployments. This evidence suggests that the localization results could be improved by designing hybrid localization schemes capable of exploiting both connectivity and RSS measurements. In the following section, this idea is explored by proposing a new SOM variant that exploits the two types of information during the training phase of the map.

5.2.1 The SOM-R algorithm

The negative effects of large connectivity values on the SOM schemes are easily understood by recalling the update rule (3.2) discussed in Section 3.1. When the network is highly connected, a large number of neurons will be within the same distance from the BMU. Therefore, the weight updates will be similar for many nodes, and the map will not be able to accurately represent the input distribution. This problem is evident in the motivating example of Section 4.1.1, where most of the estimated positions collapse toward the center of sampling space (see Figure 4.1d at page 52).

The SOM-R algorithm avoids the shortcomings of range-free localization in dense networks by using the RSS values to redefine the hop count distance d_{hop} . The idea is to augment the proximity information by sorting the one-hop neighbors on the basis of their received power. In particular, the attenuation of the RF signal between a pair of nodes i and j is measured by the *path loss* P_L :

$$P_L(i, j) = P_0 - P_{ij}, \quad (5.2)$$

where P_0 and P_{ij} have the same meaning defined in (4.3) and (4.4). Similarly to P_{ij} , P_L is also a random variable with normal distribution; the expected value for P_L is approximately zero for nodes whose separation distance is d_0 , and it increases for nodes that are far apart. It can be assumed that d_0 is small compared to the separation distances between the network nodes, therefore $P_L(i, j) \geq 0, \forall i, j$. If the network is implemented using transceivers with a typical *sensitivity* P_s ¹, we expect the path loss to increase up to a maximum value $P_{L\text{-MAX}} = P_0 - P_s$. When such value is reached, the RF power at the receiver will equal P_s and a further increment in P_L will cause the communication to fail with a high probability. Based on these considerations, the new neighborhood function used in the SOM training algorithm is:

$$h_{cj}^{(\text{PL})} = \exp\left(-\frac{d_{\text{hop}}^{(\text{PL})}(c, j)^2}{2\sigma^2}\right). \quad (5.3)$$

In the expression above, $d_{\text{hop}}^{(\text{PL})}(c, j)$ uses the path-loss values to measure the distance between the BMU at index

¹The radio sensitivity is the minimum signal power that the transceiver is able to demodulate with high probability. For example, transceivers compliant with the IEEE 802.15.4 should be able to ensure a Packet Error Rate (PER) less than 1% for signal with power equal to -86dBm or above.

c and the a generic node/neurons at index j :

$$d_{\text{hop}}^{(\text{PL})}(c, j) = \begin{cases} P_L(c, j)/P_{L\text{-MAX}} & \text{if } d_{\text{hop}}(c, j) = 1 \\ d_{\text{hop}}(c, j) & \text{if } d_{\text{hop}}(c, j) \neq 1. \end{cases} \quad (5.4)$$

The modified hop-count distance is unchanged for nodes that are not in the radio range of the BMU ($d_{\text{hop}}(c, j) \neq 1$), but it has increased “resolution” for one-hop neighbors. Different nodes within the radio range of the BMU are treated differently depending on their path loss value. A $P_L(c, j)$ value close to zero will result in a small hop-count value ($d_{\text{hop}}^{(\text{PL})} \approx 0$), which in turn will cause a strong interaction ($h_{c,j}^{(\text{PL})} \approx 1$) between the BMU and node j . On the other hand, when $P_L(i, j)$ is close to $P_{L\text{-MAX}}$, the two nodes will be treated as regular one-hop neighbors.

As mentioned in Section 3.1.1, the shape of the neighborhood function, and consequently the choice of d_{hop} , is not a critical factor in the SOM learning algorithm. In using the P_L values as shown above, the intent is not to accurately model the distance between nodes, but simply to provide a mechanism to differentiate between neurons that otherwise would be at the same map distance from the BMU. Figure 5.3a shows an example of a neighborhood function based on $d_{\text{hop}}^{(\text{PL})}$.

The use of the $d_{\text{hop}}^{(\text{PL})}$ in place of the regular hop distance is the major difference between SOM-R and the other versions of the scheme. The modified neighborhood function (5.3) can be used both with SOM-V and SOM-A, but in the rest of this work only the anchored version is evaluated. The SOM-R version is similar to Algorithm 1 in Section 3.2.3 with the exception of the new map distance. Also, experimental results suggest that when the modified map distance is used, it is beneficial to train the map with a larger learning factor, therefore $\eta_{\text{max}} = 0.5$ is used instead than $\eta_{\text{max}} = 0.1$.

The SOM-R algorithm shares some similitude with other range-free schemes where RSS values have been used to complement connectivity information. For example, in the PRI scheme proposed by Li et al. [99], the received signal strength is used to compute “sub-hops” by sorting the one-hop neighbors, and Liu et al. [102] compute the node positions as intersection of concentric rings derived from the RSS. The SOM-R scheme also shares some resemblance with the work of Nguyen et al. [124], who used RSS value collected between the nodes to train a kernel-based classifier. However, while the classifier only detects if a node is contained in a given region or not, the SOM technique implements a more straightforward approach to localization that produces explicit position estimates as a consequence of the training phase of the map.

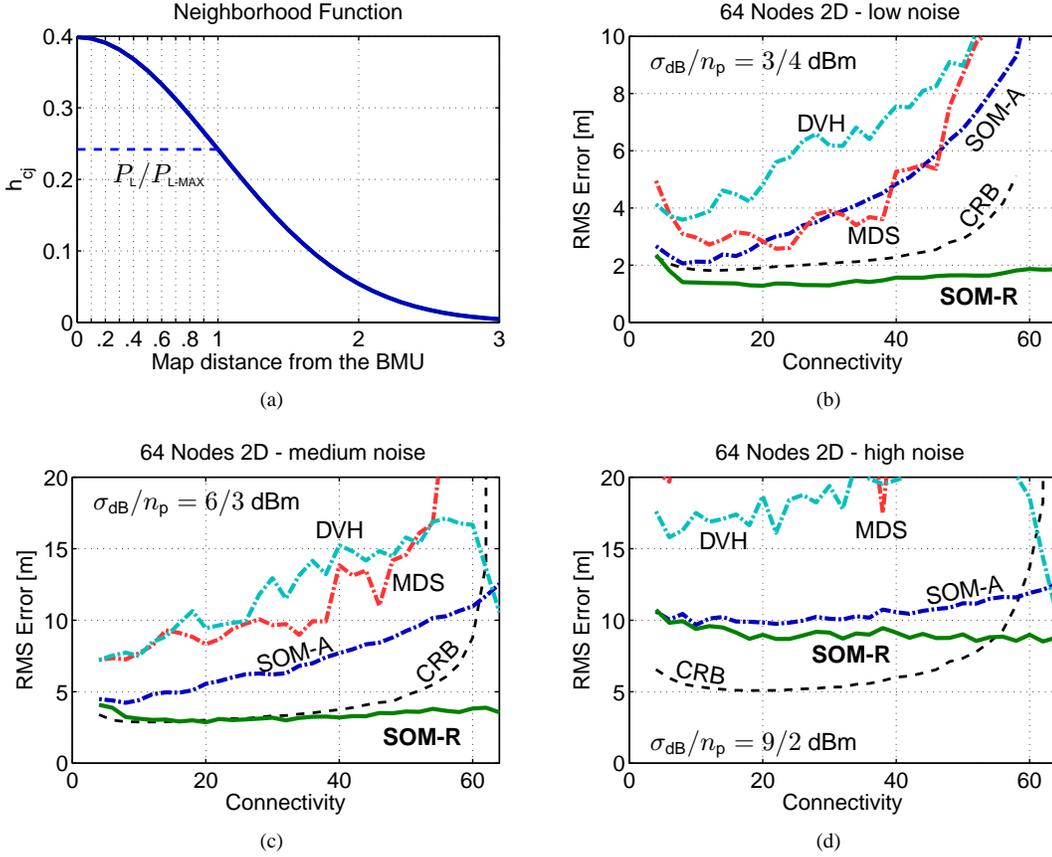


Figure 5.3: Average RMS error achieved by SOM-A, SOM-R, MDS and DVH.

5.2.2 Simulation Results

Similarly to the previous section, the RMS error is computed over 50 localization experiments on a 64-node network with four anchor nodes. Once again, the simulation accounts for different radio environments by considering the three noise levels previously used. Also, in this simulation set, the value P_{th} is used in place of the sensitivity P_s in the term P_{L-MAX} in (5.3). In general, if a threshold is not used and the value of P_s is unknown, the maximum path loss value measured within the network can be used in place of P_{L-MAX} .

Figure 5.3b,c,d show the simulation results. SOM-R produces remarkable improvements in accuracy, especially for higher values of network connectivity. In simulations with low and medium noise, and for connectivity equal to 60, the RMS error is respectively 85% and 65% lower than the value produced by SOM-A. In networks with high noise ($\sigma_{dB}/n_p = 9/2$ dBm), the RMS's error is about 25% lower than that of SOM-A. Notably, SOM-R maintains the accuracy of SOM-A for sparse networks, and it produces meaningful localization results even in fully connected networks, achieving results that are practically independent from the network connectivity.

In Figure 5.3, the SOM-R's error is sometimes lower than the CRB. This is not in contradiction with the definition of the CRB, because SOM-R uses not only connectivity constraints, but also the RSS information. However, the use of raw RSS data in the SOM algorithm is different from other approaches that use signal strength to estimate the inter-node distances. To produce such range estimates, for example using the MLE (see Section 4.3.1), knowledge of the propagation model parameters n_p is required², which in turns involves collection of a large set of controlled measurements and adds to the costs of the localization service implemented. The SOM-R scheme can localize nodes deployment in environments for which the parameters of the propagation model are unknown.

Localization Results in Larger Networks

An additional simulation set considers networks similar to the ones in previous sections, but with a larger number of nodes. Figure 5.4a,b,c,d show the average RMS error and the CRB computed for networks with 200 nodes deployed in a square region with 400 m sides. The ratio $\sigma_{dB}/n_p \in \{3/4, 6/3\}$ used in each simulation set is reported below the error plots.

When the SOM algorithm is executed with 5000 iterations, its localization error still compare favorably to that of MDS and DVH, but the values are not as close to the CRB as in the simulations reported in the previous section (see Figures 5.4a,b). The results can be improved by increasing the number of iterations. As shown in Figures 5.4c,d, training the map with 20000 samples improves the localization error and reduces the differences with the CRB. Similar results are achieved on 400 node networks deployed in a square region with side measuring 1000 m (see Figures 5.4e,f). In this case, the SOM map is trained with 50000 samples from an uniform distribution. The execution time necessary to run the MATLAB algorithm on a 2.66 GHz desktop computer is about 17 seconds.

As previously discussed, the SOM approach allows a system designer to trade accuracy for execution time. In large deployments, it is reasonable to assume that some of the devices will be powerful enough to support extensive computation. In this case, the localization results can be improved by increasing the number of iterations. Figure 5.5 shows the localization as a function of the number of iterations for three topologies with 100, 200, and 400 nodes. These plots are similar to the ones in Figure 3.10 but consider larger networks. In conclusion, different from other neural network techniques, there is no risk of overtraining the map when using a large number of training samples.

²The basic MLE estimator only requires knowledge of the the parameter n_p (in addition to P_0 and d_0). To compute an unbiased version of the same estimator, knowledge of the parameter σ_{dB} is also necessary [136].

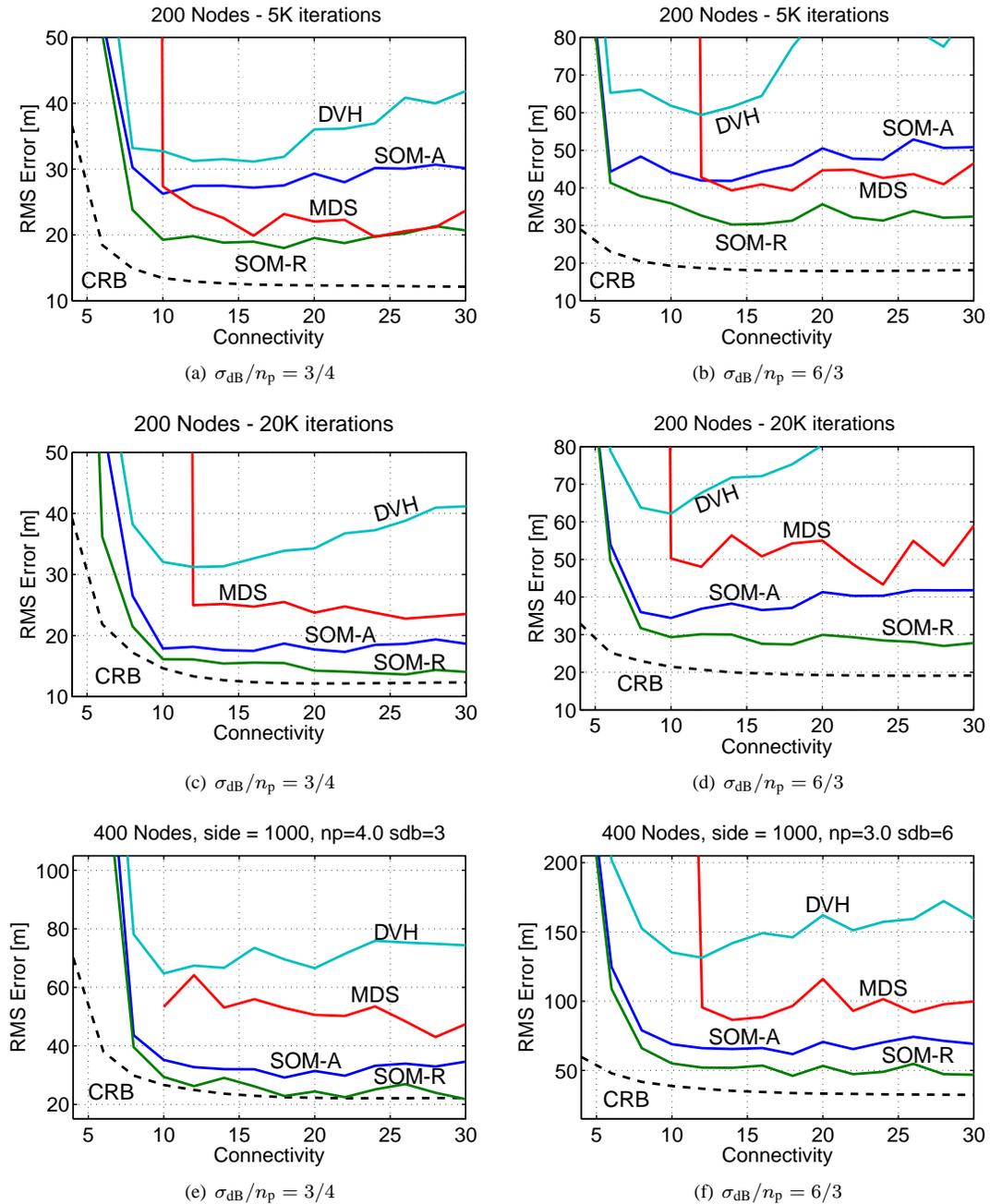


Figure 5.4: Simulation results for 200 and 400 node deployment for different connectivity values.

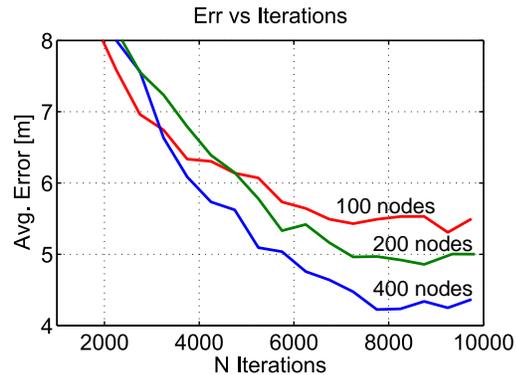


Figure 5.5: Average localization error computed using SOM for an increasing number of iterations.

5.2.3 Localization Using RSS Data From Real Deployments

This section extends the evaluation of SOM-R and the other SOM variants by presenting localization results obtained from data measured in some ad-hoc networks.

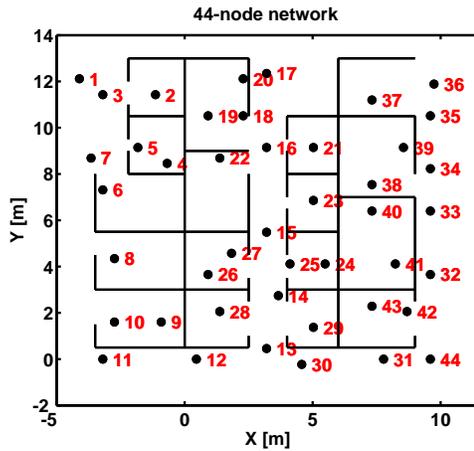
Test Case 1: 44 Node Network, Medium Noise

The first test case uses the SOM-R scheme to localize the nodes in Figure 5.6. This is the same network used in the motivating example in Section 4.1.1. The nodes in this network are all in the radio range of each other and the estimated parameters for propagation model are $\sigma_{dB} = 3.91\text{dBm}$, and $n_p = 2.3$ [136]. The ratio σ_{dB}/n_p is equal to 1.7 dBm, close to the value used to simulate networks with medium noise ($\sigma_{dB}/n_p = 6/3$ dBm).

The SOM-R scheme is executed with 5000 samples from a uniform distribution computed using (3.8). The localization experiment is repeated fifty times, changing the seed of the random number generator each time. The average RMS error achieved by the SOM-R is equal to 2.167 m, with the best and worst localization attempts that produced an error of 1.875 m and 2.518 m respectively. Comparison of the SOM-R's error with previously published results shows that SOM-R achieves performance similar to those reported by Patwari et al. [136] for a centralized MLE estimation scheme (see Figure 5.6). Note that SOM-R's results are obtained without knowledge of the parameters n_p and σ_{dB} .

Test Case 2: 46 Node Network - High Noise

The second test case uses data from a 46 node network deployed in an indoor space measuring approximately 14 m \times 10 m. The nodes use a 2.4 GHz transceiver and are arranged in a grid as shown in Figure 5.7a. Some grid positions are missing due to node malfunctioning at the moment of the test.



Scheme	RMSE	Reference
MLE	2.18 m	[136]
MDS	4.30 m	[40]
dwMDS	2.48 m	[40]
SOM-R	2.17 m	this chapter

Notes

- Nodes 3,10,35 and 44 are used as anchors in the localization process.
- The MDS schemes evaluated by Costa et al. [40] uses range estimates from RSS values.

Figure 5.6: The 44-node network with RSS measurements described by Patwari et al. [136] and published localization result for the same network.

The RSS data were collected by exchanging 100 messages between each pair of nodes and computing the average of the collected values. The deployment area, an empty office space with some metallic fixtures, was relatively uncluttered. Despite the LOS communication between most pair of nodes, the RSS data manifest a significant level of variability, presumably due to multi-path reflection from the metallic walls on the perimeter of the area and due to different antenna orientation. In fact, the nodes were equipped with an integrated antenna and were randomly oriented. Using the measured data, the propagation model parameters were found to be equal to $\sigma_{dB} = 8.13$ dBm and $n_p = 2.74$, resulting in a ratio close to the one used to simulate noisy networks. In absence of previously published results, the performance of SOM-A and SOM-R are evaluated by comparing the localization error against that of MDS and DVH. Figure 5.7b shows the localization results for different connectivity levels that were obtained by varying the threshold P_{th} . The results are qualitatively similar to those obtained in simulating networks with high ratio σ_{dB}/n_p .

Test Case 3: 38 Node Network in a 3D Space - High Noise

The last test case is based on RSS measurements from a 38 node network deployed in an indoor 3D space (see Figure 5.7c). The data is freely available on the ENALAB web site³ and the measurements are discussed in detail by Lymberopoulos et al. [110]. Similarly to the previous case, the authors found different antenna orientations and multi-path to be source of significant variability in the RSS data, which exhibit a low correlation with the distance. Figure 5.7d shows the localization error for the four scheme considered. Again, the error of SOM-A and SOM-R is significantly lower than the error of DV-HOP and MDS.

³http://www.eng.yale.edu/enalab/XYZ/data_set_1.htm

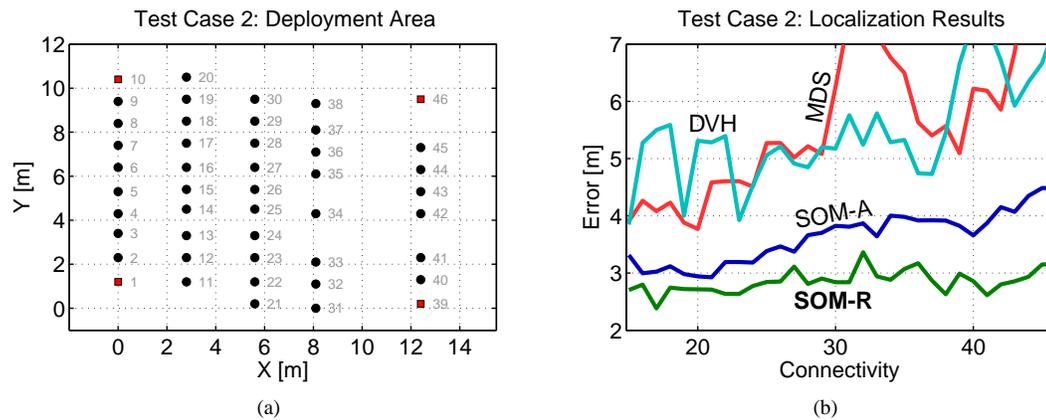


Figure 5.7: Node deployments and localization results for test cases 2. Red squares represent the anchor nodes used.

5.3 Localization in Anisotropic Deployments

Anisotropic layouts result from deploying sensors in regions with obstacles (e.g. tall buildings), or when localized node failures lead to “holes” in otherwise isotropic topologies. It is known that localization in anisotropic networks is challenging for schemes that use the hop count values as an approximation of the true node distance (e.g. MDS and DV-HOPS). In fact, while this approach works well when the path connecting any two nodes lies approximately on a straight line, it generates large errors in presence of obstacles. Any two nodes can be physically close even if their hop distance is large.

The large error in anisotropic networks has motivated alternative approaches. For example, some schemes use MDS to compute small local maps that are then stitched together into a global map [153, 74]. Although this approach yields to an interesting distributed scheme, the process of map stitching increases the

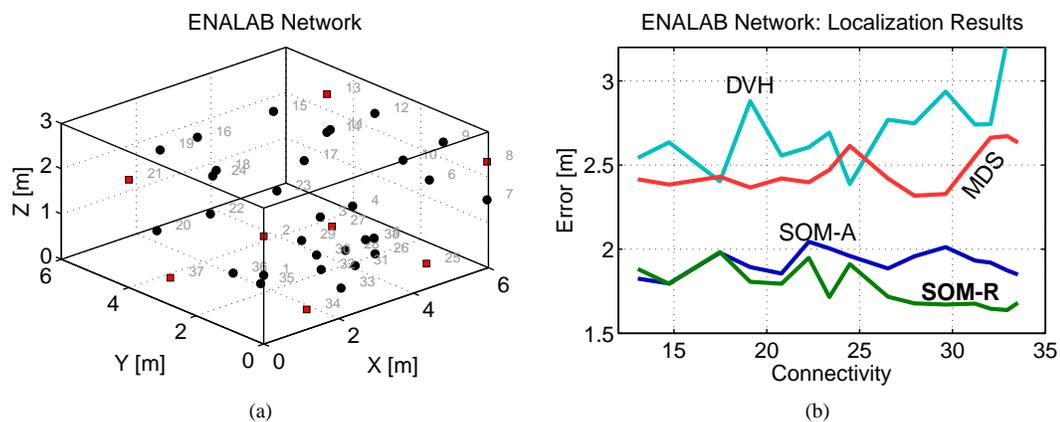


Figure 5.8: Node deployment and localization results for nodes placed in a 3D network.

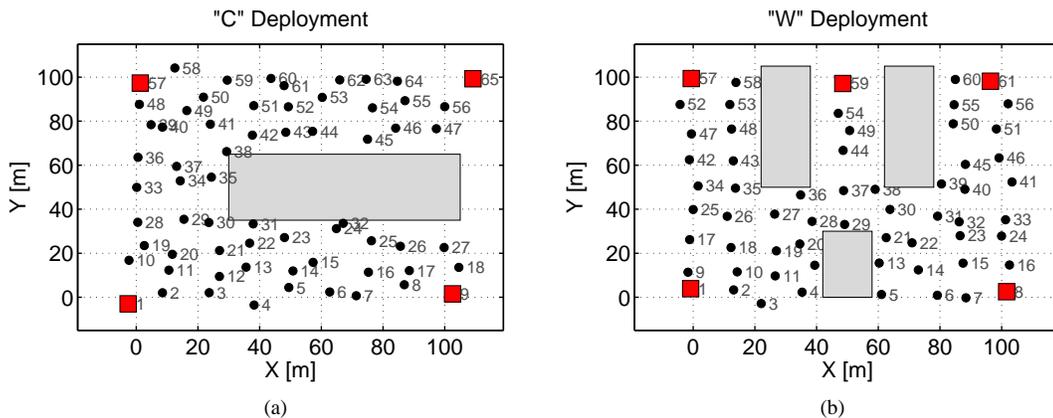


Figure 5.9: Sample anisotropic topologies. Red squares are the anchor nodes.

complexity of the solution and is susceptible to large errors when the connectivity is low. If some components of the network are not rigidly connected, the sub-maps may get stitched together with the wrong relative orientation [115].

A scheme capable of localizing irregular networks without having to partition the map and encumber the complexity of map stitching would simplify localization in practical applications. The SOM schemes is potentially well-suited for this task because its learning algorithm is designed to mainly exploit the interaction of nodes within a short hop distance, while nodes that are several hops away have a weak interaction and do not directly influence each other.

In this section, the performance of the proposed schemes are evaluated by generating simulation scenarios with few large obstacles blocking the communication between nodes. Two sample topologies are shown in Figures 5.9a and 5.9b. The figures refer to these two simulation scenarios as “C” and “W” deployments. For each topology, 50 networks were generated using the same noisy grid model described in Section 3.3.1. The only difference is that nodes are not allowed in correspondence of the obstacles. The connectivity information are obtained by first sampling pairs of RSS values using the shadowing model of Section 4.1.5 with parameters $n_p = 4$, $\sigma_{dB} = 3\text{dBm}$, and then selecting a threshold as in (4.5). Two different threshold values were used, resulting in half the networks having connectivity around 6.5 and the other half above 12. For each network, localization was repeated 25 times using different realization of the RSS values; the results of these repetitions were used to compute the RMS error and compare it with the CRB.

Table 5.1 shows the simulation results; in addition to the average RMS error, the table also reports the average localization error relative to the communication range R . The maximum range was computed using the parameter of the shadowing model and the threshold, so it should be intended in the sense of *expected*

	"C" conn = 6.68		"C" conn = 12.49		"W" conn = 6.69		"W" conn = 12.92	
Scheme	Err(R)	RMS	Err(R)	RMS	Err(R)	RMS	Err(R)	RMS
MDS	1.44	37.0 m	1.12	45.7 m	1.15	28.8 m	0.69	27.2 m
DVH	0.86	20.3 m	0.63	22.1 m	0.74	18.7 m	0.53	21.0 m
SOM-A	0.31	7.7 m	0.27	9.5 m	0.31	7.9 m	0.24	9.9 m
SOM-R	0.29	7.3 m	0.21	7.8 m	0.28	7.3 m	0.18	7.4 m
CRB		6.8 m		7.0 m		7.8 m		6.5 m

Table 5.1: Localization results in anisotropic networks.

maximum communication range. The SOM-A's and SOM-R's results were computed using the same training distribution described in Section 3.5.1 (i.e. without knowledge of the obstacles' presence.)

The results shows that SOM-A achieves an average error comparable to that of uniform networks. On average, the SOM-A's error is 75% and 60% lower than the error of MDS and DVH respectively. The errors for MDS and DVH on the "C" topology are comparable with previously published results reported by Vivekanandan and Wong [167] and Niculescu and Nath [125] for the same type of network. SOM-R generates an additional 15% error reduction with respect to SOM-A.

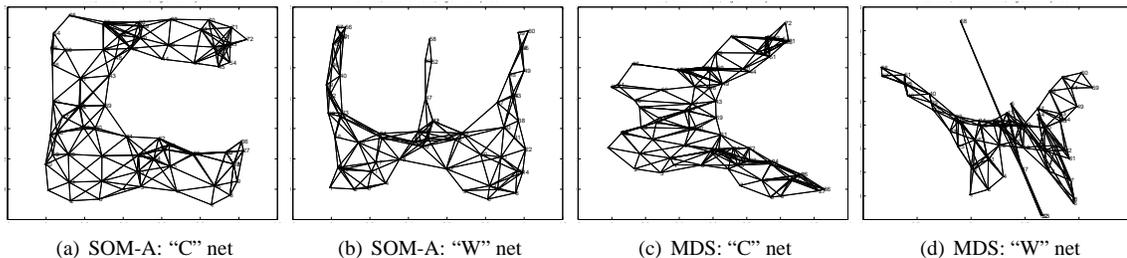


Figure 5.10: Sample results for anisotropic layouts: the SOM-A algorithm reduces the average localization error of 75% with respect to MDS.

Chapter 6

Localization using Directional Antennas

The directional antenna (DA) is an established technology that is effective in improving the performance of wireless networks. The ability to radiate the RF signal toward the receiver results in a more efficient utilization of power, in a better link quality, and in an increased transmission range. In addition, since communication is restricted in space, interferences between devices are reduced and *spatial reusability* can be exploited to increase network capacity and throughput [69, 177, 18].

The improved performance of DAs make them suitable for cellular towers and base stations, but their use in ad-hoc wireless networks is not equally widespread. One complication is the need for specific protocols capable of supporting directional communication. Many extensions to the popular IEEE 802.11 MAC layer have been proposed in the literature [18, 85, 37, 162], and several other works have addressed the problem of directional routing [156, 169, 36, 76]. But despite the research effort produced, the lack of central coordination typical of ad-hoc networks makes it difficult to fully take advantage of the directive technology, especially when nodes are mobile [16].

Similarly, DAs have been considered not suitable for sensor network applications. In WSNs, simplified¹ communication protocols can be adopted [147, 44], but the complexity of the DA technology (mainly cost and size) seems to contrast with the need to keep the nodes simple, small and inexpensive. Nevertheless, as radio communication moves to higher frequencies and antenna dimensions shrink, the use of DAs on sensor nodes appears not only feasible [96], but also desirable to compensate for the higher path loss intrinsic of shorter wavelengths, to ensure higher link quality, and to implement a form of antenna diversity [176].

¹The design of MAC protocols using directional antennas is in part simplified by the fact that sensor nodes are static and transmissions are sporadic. In addition, directive routing protocols can exploit the fact that sensor nodes typically transmit all their data to a single aggregation point (a cluster head or a base station).

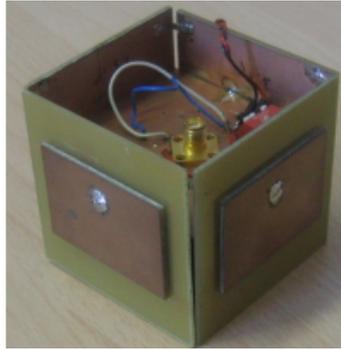


Figure 6.1: One of the switched-beam antennas developed in collaboration with Microelectronics Lab at the Università Degli Studi di Firenze, Florence, Italy.

Another advantage of DAs is that they can be used to estimate the angular position between pairs of nodes, and this information can be used to implement localization schemes based on *Angle of Arrival* (AOA) information [121, 175, 111]. This approach improves over RSS ranging schemes for two reasons:

1. AOA estimates can be obtained without assumptions on the propagation model that relates the RSS to the distance. As shown in the algorithms in Section 6.1.3 and 6.2, only knowledge of the radiation pattern is required to estimate the AOA of the incoming messages. On the other hand, RSS ranging is based on the propagation model for the RF signal in a given environment (see Section 4.3.1). Extensive measurement campaigns are needed in order to identify a suitable propagation model and estimate its parameters.
2. AOA-based localization requires a lower number of anchor nodes than localization based on distance estimates. In the 2D case, only two anchors are needed when using angles, while at least three reference nodes are needed when using distance estimates (see Figure 2.6).

The last part of the research work presented in this dissertation has been dedicated to the study of DAs and evaluation of AOA estimation algorithms. As a result of joint work with the Microelectronics Lab, Università Degli Studi di Firenze, ITALY, two different antennas suitable for AOA estimation have been designed, prototyped and tested. The use of these antennas and their application to localization is described in the following sections. Additionally, the use of AOA information in the SOM localization scheme is described in Section 6.3

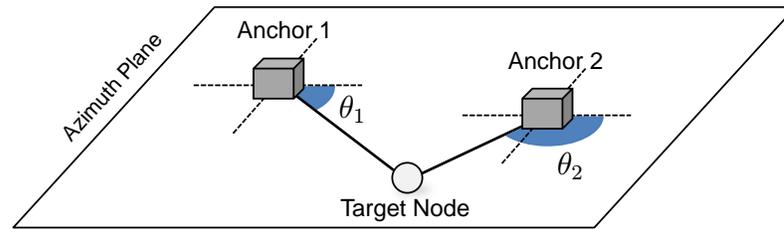


Figure 6.2: Azimuth AOA estimation.

6.1 Azimuthal Angle of Arrival Estimation

Several examples of localization schemes that use angular information have been proposed in the literature. These solutions use DAs to estimate the AOA of messages transmitted by other nodes located in the same plane of the antenna. This configuration enables estimation of the azimuthal angle θ (see Figure 6.2); absolute localization in 2D can be achieved by combining angle estimates from two reference points.

This section describes a directional antenna suitable for azimuthal AOA. An extension of this principle is discussed in Section 6.2, which proposes a directional antenna capable of estimating both the azimuth and elevation angle of the incoming messages.

6.1.1 Four Beam Directional Antenna (FBDA)

The first antenna developed in collaboration with the MicLab at Univ. of Florence is a unit dubbed *Four Beam Directional Antenna* (FBDA). The FBDA is composed of four coaxially fed planar patch antennas arranged in a “box like” structure as shown in Figure 6.1. Each face is realized on a two-layer RF4 substrate [15] having planar dimension of $56 \text{ mm} \times 56 \text{ mm}$ and thickness of 2.4 mm. The four patches, which operate in linear polarization, share a common design that has been optimized using the Ansoft-HFSS CAD [9] to work in the 2.4 GHz ISM band. The mechanical arrangement of the four patches and their coaxial feeding is such that the vertical axis of the box coincides with the intersection of the E-planes of the single patches (i.e. the E-field is perpendicular to the ground).

The RF signal is distributed to the four faces by a single-pole four-trough switch, which is controlled by two digital lines and allows the host node to dynamically select the face to use. The losses due to the switch, the distribution network and the mismatches are about 1.5 dB within the selected ISM band.

The characterization in the anechoic chamber has given the patterns reported in Figure 6.3. In spite of the low-cost substrate and reduced thickness, the patch gains measured at the external SMA connector,

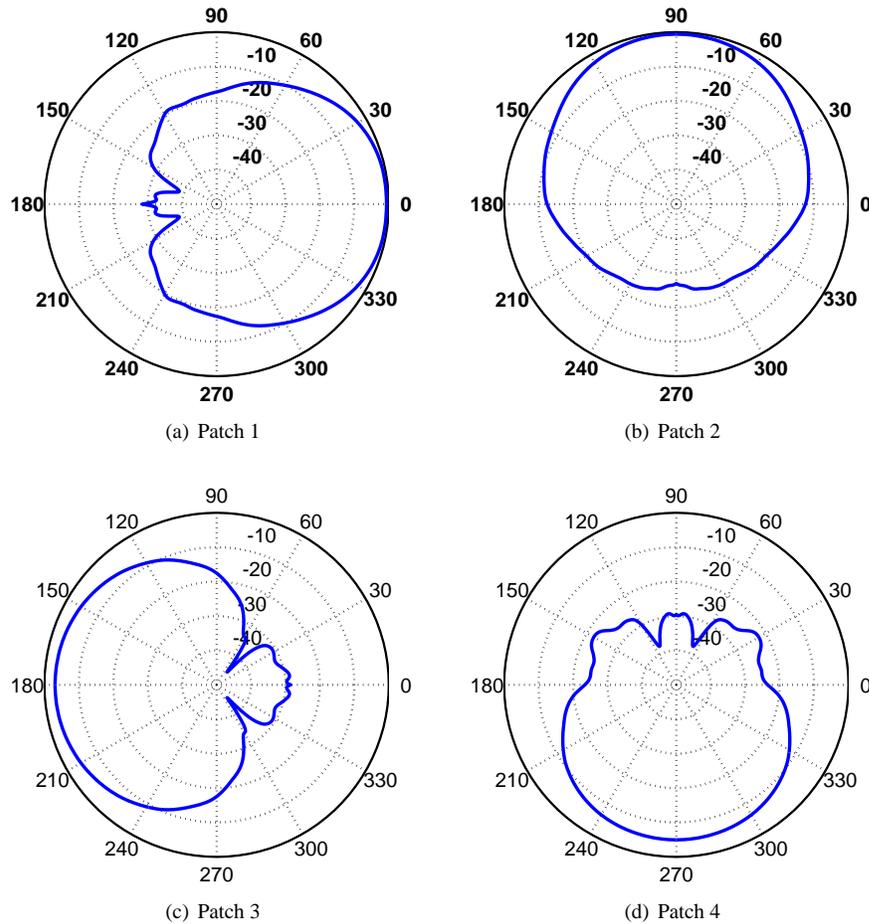


Figure 6.3: Radiation patterns of the four antenna faces. The patch used for transmission/reception is selected using two digital lines.

hence including the losses listed before, are comprised between 8.3 dBi and 7.5 dBi. Figure 6.3 shows that the combined patterns ensure an approximately uniform coverage of the 360° horizon.

6.1.2 Principle Of Operations

Consider a target node equipped with an omni-directional antenna. The target node transmits radio packets to a base station equipped with a FBDA (see Figure 6.4). The goal is to estimate the angle of arrival of such messages relative to a reference system aligned with the antenna's axes. Measuring the RSS on two antenna faces provides sufficient information for AOA estimation [11, 175, 111]. According to the Friis' free space

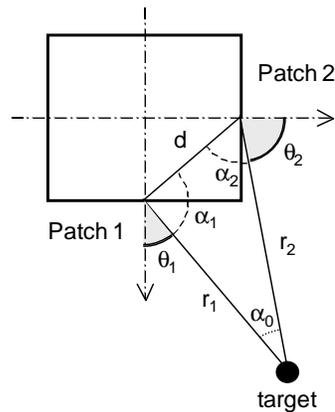


Figure 6.4: Distances and angles of a target node relative to patch 1 and 2.

equation, the received power P_{r1} and P_{r2} on faces 1 and 2 is:

$$P_{r1} = \frac{P_t G_t G_{r1}(\theta_1)}{r_1^2} \left(\frac{\lambda}{4\pi} \right)^2 \quad (6.1)$$

$$P_{r2} = \frac{P_t G_t G_{r2}(\theta_2)}{r_2^2} \left(\frac{\lambda}{4\pi} \right)^2, \quad (6.2)$$

where P_t is the power of the target's messages transmitted using an omnidirectional (OD) antenna with gain G_t . The values $G_{r1}(\cdot)$ and $G_{r2}(\cdot)$ are the angular gains of the patches 1 and 2 and depend on the angles of arrival θ_1 , θ_2 . Since the relative angle between each patch is fixed (90° in the proposed antenna), the ratio between the power received on the two patches is equal to the ratio between the antenna gains for angles θ_1 and $(\theta_1 - \pi/2)$:

$$\frac{P_{r1}}{P_{r2}} = \frac{G_{r1}(\theta_1)}{G_{r2}(\theta_1 - \pi/2)}. \quad (6.3)$$

The relation above holds when the distance of the target from the directional antenna is much larger than the distance between the faces of the antenna itself. If $r_1, r_2 \gg d$, then $r_1 \cong r_2$ and $\alpha_0 \cong 0$, see Figure 6.4. This scenario covers most cases of practical interest. Under the same assumption, the equations that relate the power received on the other patches can be derived (see Figure 6.5a,b):

$$\frac{P_{r1}}{P_{r3}} = \frac{G_{r1}(\theta_1)}{G_{r3}(\theta_1 + \pi)}, \quad (6.4)$$

$$\frac{P_{r1}}{P_{r4}} = \frac{G_{r1}(\theta_1)}{G_{r4}(\theta_1 + \pi/2)}. \quad (6.5)$$

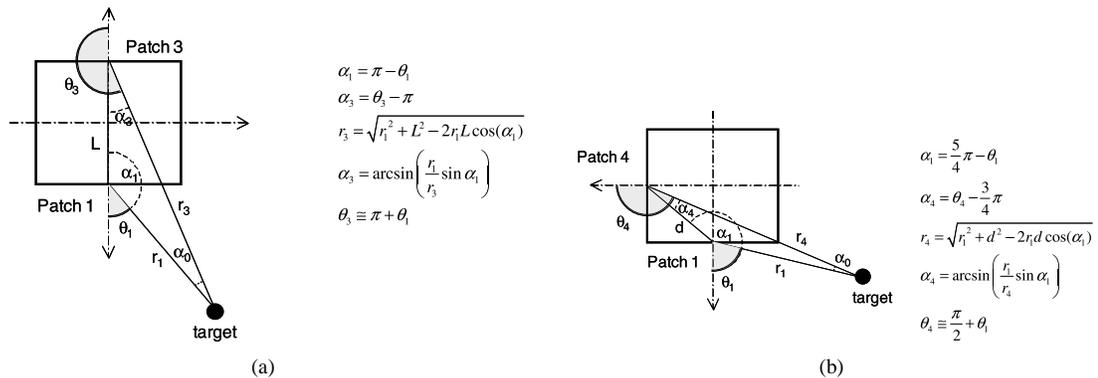


Figure 6.5: Distances and angles of a target node relative to patches 1-3, and patches 1-4.

Although measuring the RSS values on patches 3 and 4 is not strictly necessary for AOA estimation, the availability of additional measures improves the robustness of the estimation process. The angular relations above can be combined in the following system of equations:

$$DP = DG(\theta) + V, \quad (6.6)$$

where $DP = [(P_{r1} - P_{r2}) (P_{r1} - P_{r3}) (P_{r1} - P_{r4})]^t$ is a column vector containing the power differences between the patch n.1 and the other faces (in dBm), and DG contains the gain differences for any angle θ (in dBm):

$$DG(\theta) = \begin{bmatrix} G_{r1}(\theta) - G_{r2}(\theta - \pi/2) \\ G_{r1}(\theta) - G_{r3}(\theta + \pi) \\ G_{r1}(\theta) - G_{r4}(\theta + \pi/2) \end{bmatrix}. \quad (6.7)$$

Finally, the vector $V = [v_1 \ v_2 \ v_3]^t$ models the noise in the measurements and the effect of inaccurate knowledge of the radiation patterns.

6.1.3 AOA Estimation

Given (6.7), a solution to estimate θ is to use a *Least Square* (LS) estimator. The LS estimator computes the angle θ that minimizes the *least squared error norm* between the measured data and the function $DG(\theta)$ evaluated using the values from the radiation patterns:

$$\hat{\theta} = \arg \min_{\theta} \|DP - DG(\theta)\|^2. \quad (6.8)$$

Although no attempt was made to characterize the error affecting the measured data, it is known that when the noise component V in (6.6) is Gaussian with zero mean, the estimator (6.8) coincides with the maximum likelihood estimator [166]. Therefore, this solution is equivalent to previous approaches that have evaluated AOA estimation in condition of normal distribution of the noise (e.g [11]).

Multiple Signal Classification (MUSIC)

Another technique to obtain AOA estimates is based on the *Multiple Signal Classification* (MUSIC) [87] algorithm. The signal impinging on the four antenna faces can be expressed by the following relation:

$$\begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix} = \begin{bmatrix} G_1(\theta) \\ G_2(\theta) \\ G_3(\theta) \\ G_4(\theta) \end{bmatrix} s(t) + \begin{bmatrix} n_1(t) \\ n_2(t) \\ n_3(t) \\ n_4(t) \end{bmatrix}, \quad (6.9)$$

where $x(t)$ is the signal (in volts) at the output of the antenna, $s(t)$ is the signal transmitted by the target and $n_i(t)$ are *Additive White Gaussian* (AWG) noise components.

The MUSIC algorithm produces a “spectrum” $P(\theta)$ that exhibits peaks for angles θ close to the true AOA of the incoming signals. The spectrum is computed as a result of an algorithm that includes the following steps:

1. **Data Collection.** A sequence of RSS values on the four antenna faces is collected by exchanging radio messages.
2. **Covariance Estimation.** The *Spatial Covariance Matrix* R is estimated using the sequence of available RSS values.
3. **Singular Value Decomposition.** The matrix R is decomposed using *Singular Value Decomposition* (SVD). Assuming a single signal source (the target node), one of the eigenvectors is related to the target’s messages, while the other three are related to the noise.
4. **Projection.** The *steering vector* $G(\theta) = [G_1(\theta), \dots, G_4(\theta)]$ is projected onto the subspace spanned by the noise eigenvectors:

$$P(\theta) = \frac{G(\theta)^H G(\theta)}{G(\theta)^H \hat{\Pi}^\perp G(\theta)}, \quad (6.10)$$

where the matrix $\hat{\Pi}^\perp$ contains the three noise eigenvectors.

```

function [Pm] = DOA_music(G,X,M)
% AOA estimate using music algorithm.
% G: Steering Vector - [360 x 4] matrix
% X: RSS samples - [4 x n] matrix
% M: number of signal sources (typ. M = 1)

L = size(X,1);
n = size(X,2);

% replace NaN values
[i,j] = find(isnan(X));
X(i,j) = -95;

% Convert RSS and Gains from dB
X = 10.^(X/10);
G = 10.^(G/10);

% Compute the Spatial Covariance Matrix
Rh = zeros(L);

for i = 1 : n
    Rh = Rh + X(:,i)*X(:,i)';
end
Rh = Rh ./ n;

% Singular value decomposition
[U,D,V] = svd(Rh);

% separate signal component from noise

% signal
Us = zeros(4);
Us(:,1:M) = U(:,1:M);
Ds = zeros(4);
Ds(:,1:M) = D(:,1:M);

% noise
Un = zeros(4);
Un(:,M+1:4) = U(:,M+1:4);
Dn = zeros(4);
Dn(:,M+1:4) = D(:,M+1:4);

% compute the music spectrum
Pm = zeros(1,length(G));
H = Un*Un';
for i = 1: length(G)
    Pm(i) = (G(i,:)*G(i,:))'/(G(i,:)*H*G(i,:))';
end

```

Figure 6.6: MATLAB Code for the implementation of the music algorithm

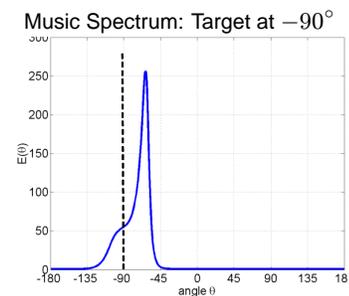
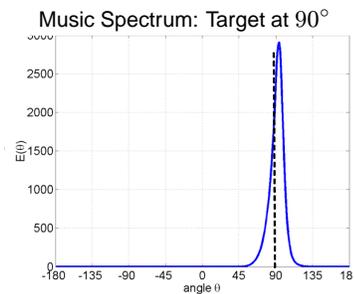
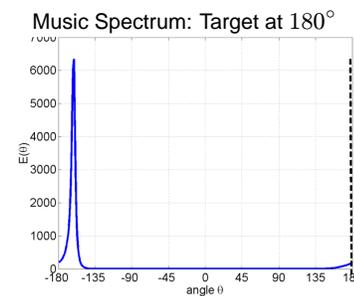
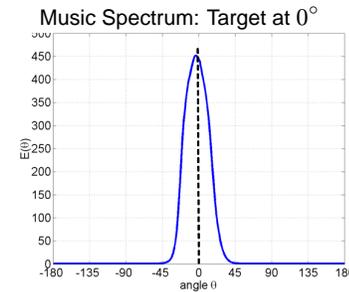


Figure 6.7: Music spectrum computed from measured data for the target at 0°, 90°, 180° and 270°.

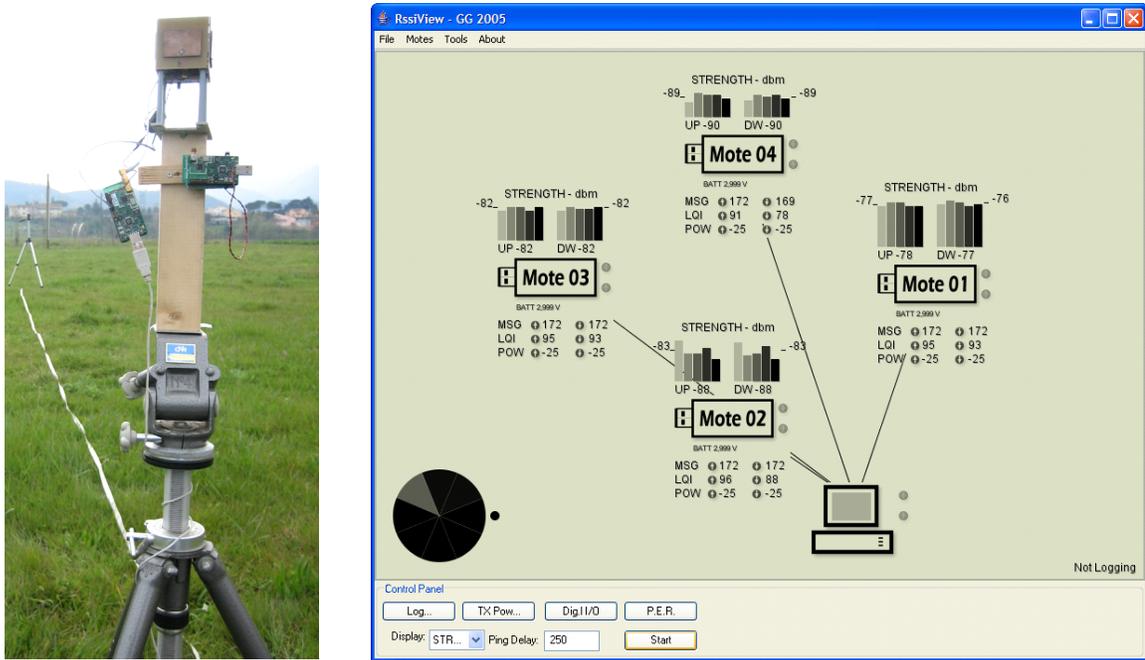
Figure 6.6 contains the MATLAB code used to implement the MUSIC algorithm. Figure 6.7 shows examples of the MUSIC spectrum computed using RSS values measured in a open field with the target positioned at 0° , 90° , 180° and 270° with respect to the directional antenna. In the four cases analyzed, the estimated AOA, which coincides with the position of the peak, is close to the true angular position of the target.

6.1.4 AOA Estimation Results using In-Field Measurement Data

Data collected during in-field experiments were used to evaluate the error of the AOA estimation algorithms described in the previous sections. Four sets of measurements were collected by placing the target node at about 3 meters from the base station, with both nodes elevated of 1.2 meter above the ground. The data was collected in two different scenarios: an open-field (see Figure 6.8a) for the first set of measurements, and a location in proximity of a building in the second case. During the experiments, the node with the directional antenna was rotated around its vertical axis on 24 different angular positions, spaced by 15° each. The RSS values on each patch were collected by transmitting bursts of 50 data packets from the target node, and each experiment was repeated using two different level of transmission power (-25 dBm and -15 dBm).

Figure 6.8.b shows the estimation error of the LS and MUSIC algorithms for the two cases: the columns marked with an “A” refer to the measurements in the open field, while the columns marked with “B” refer to the experiments made in proximity of the building. The average error is comprised between 9.8° and 23° , with noticeably larger values in the second set of experiments. In that case, reflections of the signal from nearby building are a source of noise that degrades the accuracy of the estimation process.

The radiation patterns used in the two algorithms were measured for an antenna that was different from the unit used during the tests. Since the antennas are hand-built and accurate control of the antenna characteristics is not possible, part of the error originates from imperfect knowledge of the radiation patterns of the antenna used. To mitigate these errors, AOA estimation was repeated considering radiation patterns computed using the measurements made at -25 dBm. The error achieved in this case is sensibly lower than the previous case, with values that are comprised between 5° and 13.25° (See Figure 6.8.c).



(a) A view of the setup used to collect the RSS value to test the Angle of Arrival estimation algorithms.

	A (-25dBm)	A (-15dBm)	B (-25dBm)	B (-15dBm)
LS	10.4 (7.9)	13.6 (12.21)	15.6 (15.2)	23.0 (23.8)
MUSIC	9.8 (7.6)	10.6 (9.8)	17.1 (17.5)	15.6 (17.0)

(b) Results from the experiments. The value in the table represent the average error in angle estimation. The value in parenthesis are the standard deviation of the error.

	A (-25dBm)	A (-15dBm)	B (-25dBm)	B (-15dBm)
LS	–	5.1 (5.0)	–	13.25(15.5)
MUSIC	–	6.0 (8.7)	–	9.1(11.5)

(c) Average error (and std. dev. in parenthesis) when a set of measurements is used to estimate the radiation pattern of the antenna.

Figure 6.8: Result of AOA estimation experiments: the table b and c report the estimation error in two cases (A and B) and using two transmission levels (–25 dBm and –15 dBm).

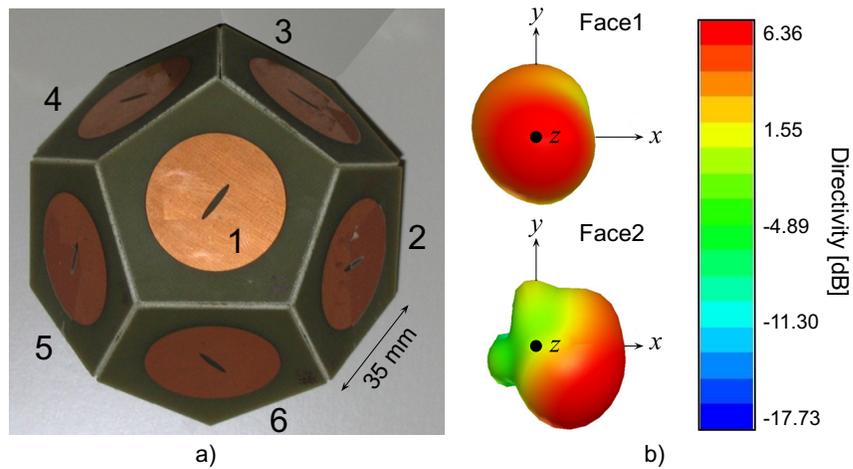


Figure 6.9: a) Bottom-up view of the switched-beam directional antenna used to implement the *single-anchor* localization system. b) Simulated radiation patterns when faces 1 and 2 are active.

6.2 Indoor localization Using a Single Anchor Node

This section describes a localization system that takes the directional approach one step further by using the second antenna (see Figure 6.9) developed in collaboration with the MicLab, Univ. Of Florence, Italy. The proposed antenna implements single-anchor node localization system. The single anchor node, which serves as a Base Station (BS), is installed on the ceiling of any large indoor space, in a position unobtrusive to the users. Due to the 3D arrangement of the antenna elements, the system can locate a target by estimating both the azimuth θ_t and elevation ϕ_t AOA of the incoming messages.

The proposed solution targets applications in large rooms or indoor open spaces where installing a network of anchors is not desirable or feasible. Possible application scenarios include low-cost deployments and ad-hoc applications (e.g. emergency response). To accommodate for different sets of requirements in terms of accuracy and cost, three localization solutions are proposed. The system supports a *range-free* (proximity), a *range-based* and a *fingerprinting* localization approach. The proposed schemes, which are evaluated using RSS traces from a real deployment, show that satisfactory localization results are possible using a single anchor node.

6.2.1 Antenna Design

The antenna mounted on the BS was designed with the goal to implement a compact, low-cost system with a steerable beam capable of selectively illuminating the space underneath the BS and collecting information useful for target localization.

The proposed solution is an incoherent array of six adjacent radiating elements, assembled to form a semi dodecahedron (see Figure 6.9a). Each element is implemented in microstrip antenna technology on a pentagonal plastic substrate and fed by a coaxial probe. The operating frequency is 2.45 GHz with a bandwidth compatible with IEEE 802.11 and IEEE 802.15.x devices such as WLAN, Bluetooth and ZigBee transceivers. Given the intended use in indoor applications, the antenna elements are implemented in circular polarization technology. This design that has proven useful in mitigating multipath effects in reflective environments [75].

A single-pole six-through RF switch is used to multiplex each radiating element. Under control of the BS, the switch connects one of the six radiators to the transceiver. The inactive faces, which are terminated on matched loads, behave as dummy loads, without significantly perturbing the radiation pattern of the active patch. Figure 6.9b shows two of the six radiation patterns simulated using the Ansoft HFSS software[9]. The directivity is typical to that of a microstrip antenna, with the main lobe pointing in the direction perpendicular to the active face.

6.2.2 Localization Application

This section describes the implementation of a *proof-of-concept* application where the antenna is used to estimate the position (x_t, y_t) of a mobile target in a large classroom containing rows of desks and chairs. The antenna was placed approximately in the center of the room, two meters above the desks, and with the face 1 pointing toward the floor. RSS traces were collected on a 6×4 grid (see Figure 6.10) by exchanging bursts of 100 messages between the target and each of the six antenna faces. The measured data was used to evaluate the performance of three different localization algorithms described in the following sections.

6.2.3 Principle of operations

Let the pair of angles (ϕ_t, θ_t) define the *Direction of Arrival* (DOA) of the target's messages. The term DOA is used in place of AOA to emphasize the difference between estimation of a single angle and the pair (ϕ_t, θ_t) . According to the Friis' equation, the power received by each face i depends on its gain $G_i(\phi_t, \theta_t)$ and the target's distance. Given the small physical dimensions of the antenna, all the faces are at about the same distance from the target. Similar to what was discussed in Section 6.1.2, the differences in received power (in dBm) between two faces i and j will only depend on their gains:

$$P_i - P_j = G_i(\phi_t, \theta_t) - G_j(\phi_t, \theta_t), \quad (6.11)$$

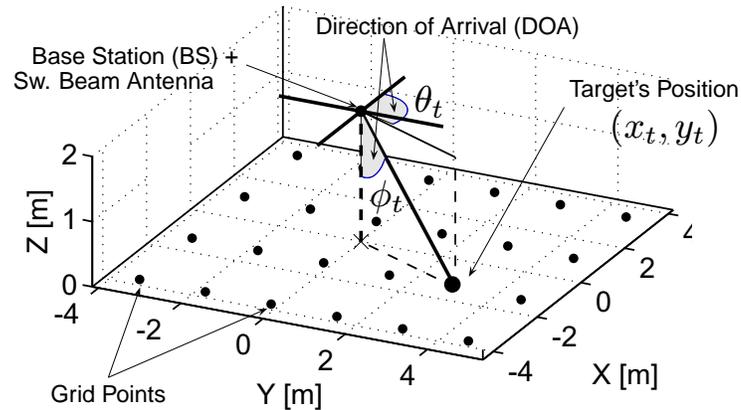


Figure 6.10: Deployment area and 6×4 measurement grid. Both the BS and the target use a TI CC2420 transceivers set to transmit at -15 dBm.

Note that with a fixed antenna position and assuming target's movements in the plane $z = 0$, there is a one-to-one correspondence between the DOA (ϕ_t, θ_t) and the target's position (x_t, y_t) . Let m and m^{-1} be the bijective functions that describe the mapping:

$$m : (x_t, y_t) \rightarrow (\phi_t, \theta_t) \quad (6.12)$$

$$m^{-1} : (\phi_t, \theta_t) \rightarrow (x_t, y_t) \quad (6.13)$$

The following sections will show how the above relations can be exploited to estimate the target's position.

6.2.4 Range-Free Localization (Proximity)

The first approach evaluated is a range-free scheme that provides coarse-grained localization. This solution bears resemblance to a simple proximity-based scheme, but instead of relying on a set of anchor nodes, it only uses measurements from the switched-beam antenna.

To implement this solution, the radiation patterns and (6.12) are used to partition the deployment area in a set of non-overlapping regions S_1, \dots, S_6 . These regions are computed by comparing the gains $G_i(m(x, y))$ seen in different locations of the deployment area (see top row of Figure 6.11):

$$S_i = \{(x, y) : G_i(m(x, y)) > G_j(m(x, y)), \forall i \neq j\}. \quad (6.14)$$

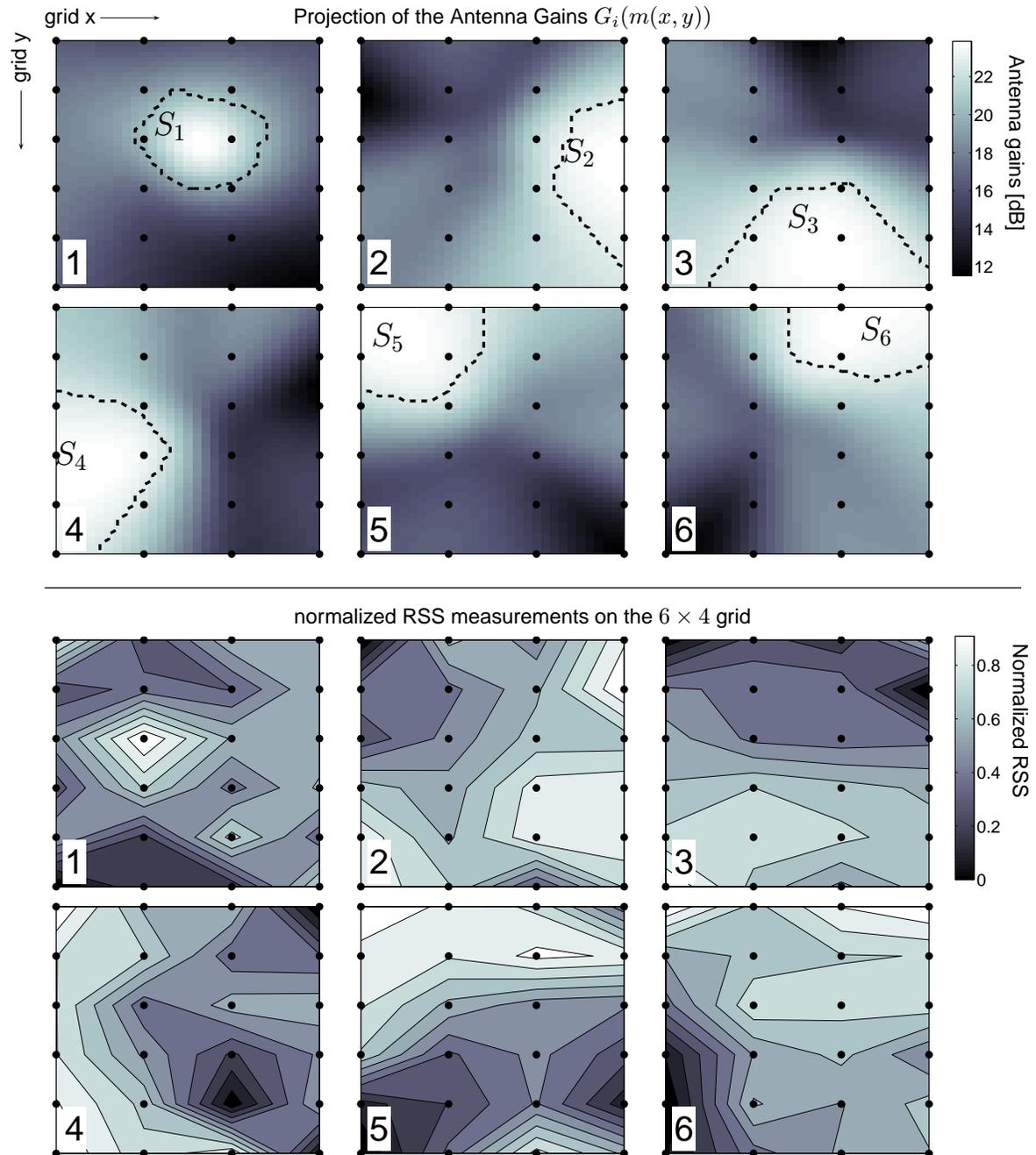


Figure 6.11: Top: antenna gains of faces 1 to 6 at different locations of the deployment area. Bottom: measured RSS values from the six antenna faces on the 6×4 grid. To facilitate comparison with the gain values, the RSS values are centered on their mean and normalized in the range $[0, 1]$.

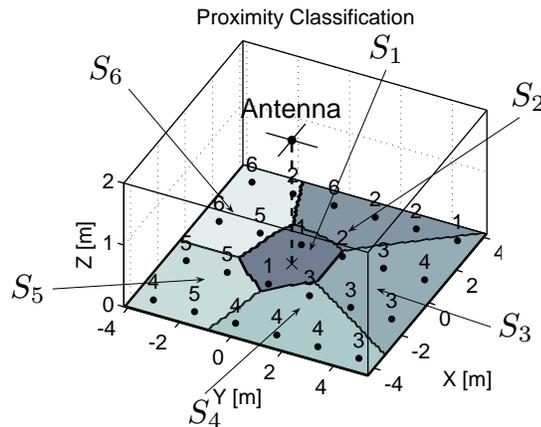


Figure 6.12: Localization results using Proximity-based classification. Average error = 2.34 m.

Each area S_i contains the locations that are best illuminated by face i . In an ideal environment, face i would receive the strongest signal (compared to the other faces) for any message sent from locations $(x_t, y_t) \in S_i$. Assuming a more realistic propagation model, the constraints on the RSS can only be formulated in a statistical sense. For example, when the signal is described by the widely adopted *log-normal shadowing model* discussed in Section 4.1.5, the average RSS (in dBm) follows a normal distribution. In this case, each area S_i contains the locations where the static expectation for the power on face i is larger than the expected values on the other faces.

Based on the previous observations, at runtime the localization algorithm assigns the target's position to the area S_i that corresponds to the face measuring the strongest (average) signal:

$$(\hat{x}_t, \hat{y}_t) \in S_{i_{\max}}, \quad \text{with } i_{\max} = \arg \max_{i \in [1,6]} \{z_i\}, \quad (6.15)$$

where z_i is the mean of the RSS values $\mathbf{z}_i = \{z_1^{(i)}, z_2^{(i)}, \dots\}$ collected by each face i at a given location. Note that the computational requirements of this approach are minimal because the regions S_i are computed off-line and only depend on the antenna position and its radiation patterns. In particular, by varying the antenna's height it is possible to adjust the size of the areas S_i and control how the deployment area is partitioned.

The measured RSS data (see bottom row of Figure 6.11) were used to evaluate the result of the assignment (6.15). Figure 6.12 shows the deployment area partitioned according to the areas S_1, \dots, S_6 and the classification results on the 6×4 grid points. Some misclassification occurred, especially between adjacent areas and in one corner of the room, where the effect of multipath was more severe. To each point it was assigned an error equal to the distance between its position and the center of the area $S_{i_{\max}}$ computed

using (6.15). The average error is equal to 2.34 m. Despite these errors, the simplicity of this approach is attractive to applications that can tolerate approximate positions. The results could be improved by computing the regions S_1, \dots, S_6 using more sophisticated models, such as the ray tracing approach adopted in a previous localization system based on a single base station [178].

6.2.5 Range-Based Localization (DOA Estimation)

The second solution implemented uses the measured RSS values to estimate the DOA $(\hat{\phi}_t, \hat{\theta}_t)$ of the incoming packets. This implementation uses the popular MUSIC approach already described in Section 6.1.3. The algorithm is similar to the azimuthal AOA estimation, but the spectrum is now evaluated for pair of angles (ϕ, θ) (See Figure 6.13). Let $P_M(\phi, \theta)$ be the spectrum produced by the MUSIC algorithm. The estimated DOA is defined by the pair of angles that yield the maximum spectrum value:

$$(\hat{\phi}_t, \hat{\theta}_t) = \arg \max_{(\phi, \theta)} \{P_M(\phi, \theta)\}. \quad (6.16)$$

Then the target position is estimated by applying (6.13):

$$(\hat{x}_t, \hat{y}_t) = m^{-1}(\hat{\phi}_t, \hat{\theta}_t). \quad (6.17)$$

Compared to the previous case, this method is computationally more expensive, but it allows for fine-grained localization. Figure 6.14a shows the localization error using the same set of RSS measurements previously shown. The average localization error is equal to 1.69 m.

This approach extends previous solutions exploiting beacon with directional antennas located on the target's plane [130, 111]. When DOA estimation is limited to the azimuth angle θ_t , single-anchor localization is not possible unless combined with distance estimates obtained from RSS measurements. In the proposed application, since both the azimuth θ_t and elevation angles ϕ_t are estimated, target positions in the $z = 0$ plane can be resolved without need of additional information.

6.2.6 Fingerprinting

The last solution tested is a fingerprinting scheme that estimates the target's position by comparing the RSS on the six antenna's faces against a database of previously measured values. This approach offers a low-computation solution that is oblivious of the RF propagation model and the antenna gains. On the downside,

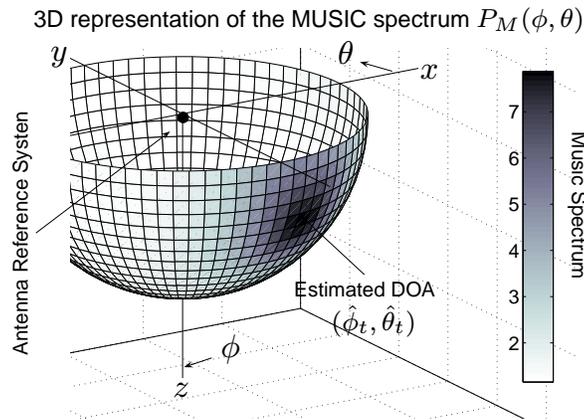


Figure 6.13: Values of the MUSIC spectrum $P_M(\phi, \theta)$ represented in spherical coordinate system centered on the antenna position. The strongest value defines the estimated direction of the incoming signal.

it requires a site survey to collect RSS signatures at several locations of the deployment area.

In the proposed solution, the database was created using the average RSS values $[z_1, \dots, z_6]$ collected on the 6×4 grid (see Figure 6.11); a second set of similar measurements was used to evaluate the localization error. Each location was estimated by first computing the Euclidean distance between the actual RSS values and the stored measurements, and then applying a *K-Nearest Neighbor* (KNN) regression algorithm [104]. The better results were achieved by setting $K = 1$, which yielded an average localization error equal to 2.32 m (see Figure 6.14b).

The average error achieved using fingerprinting is comparable to the error achieved using DOA, but it has a larger variability. The large error on some grid points was probably caused by the different type of antenna mounted on the target device in the second round of measurements.

6.2.7 Discussion

The proximity and range-based approaches discussed in Sections 6.2.4 and 6.2.5, which are based on the DOA of the target's messages, are suited for large rooms or indoor open spaces where *Line Of Sight* (LOS) communication with the BS station can be ensured. Outdoor localization can be also supported by placing the BS at a sufficient height. For example, the antenna could be installed on a tall pole placed approximately in the center of the deployment area.

When LOS communication is possible, the experimental results discussed in this chapter have shown that single-anchor 2D localization is feasible using a low-cost, RF-based system that requires zero-

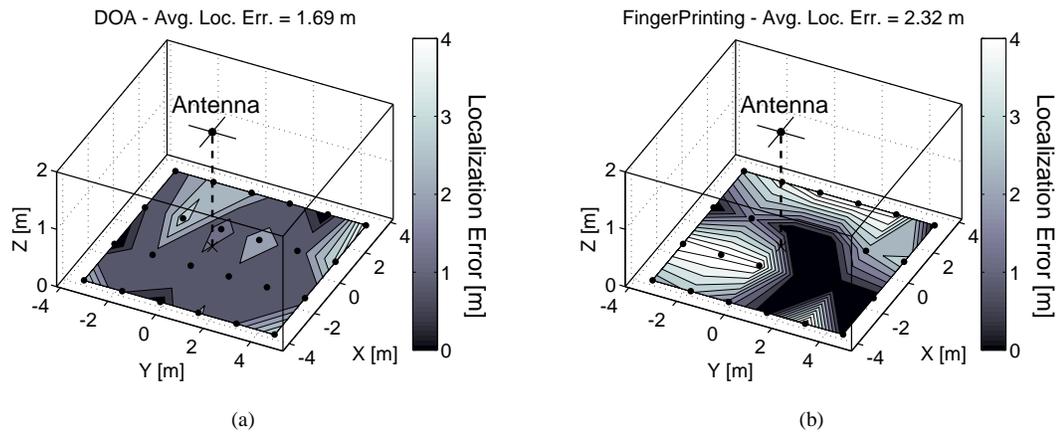


Figure 6.14: Localization results using a) DOA Estimation, b) Fingerprinting.

configurations. The reported results were obtained with an initial antenna prototype optimized for size and using simulated radiation patterns. Most likely, the error could be further reduced by using antenna faces with larger ground planes and radiation patterns measured in an anechoic chamber.

Although experiments were conducted in only one environment, larger localization errors are to be expected in cluttered environments and for non LOS communication. For such applications, the fingerprinting approach discussed in Sec. 6.2.6 represents a viable solution. The localization results are comparable to other fingerprinting solutions described in literature [104], and the use of a single anchor can alleviate the deployment costs when an infrastructure of anchors is not already available.

6.3 Use of Angle Information in Collaborative Localization Schemes

A target node at an arbitrary position can be geo-located using measurements from a single anchor node equipped with the antenna described in the previous section. Alternatively, the position of the same target can be determined by combining two AOA estimates from base stations using DA's similar to the unit described in Section 6.1. Example of localization schemes that use AOA information have been described in Section 2.3, and other solutions can be found in the literature. For example, Nasipuri and Li [121], and several other authors (e.g. [38, 130, 122, 45]) have proposed schemes where each node estimates its position by listening to directional beacons transmitted by anchor nodes in the corners of the deployment. Yang et al. [175] have evaluated the use of DA's to estimate the position of a mobile node by combining AOA and RSS-ranging measurements, and Malhotra et al. [111] have extended this approach for use of anchor nodes that are not aligned with respect to a common reference system.

As discussed in the introduction regarding GPS, localization schemes that use measurements from directional beacons represent a special instance of the localization problem. In fact, every node to be localized need to be in the radio range of the base station(s) with the directional antenna. If some of the nodes are located outside the coverage area of the BS's, their position can be computed using collaborative localization schemes. This approach requires the presence of additional nodes configured to work as a network, but it will work even if some (or most) of the units are not in the radio range of the reference nodes. Examples of collaborative schemes using AOA information are solutions based on the MLE [117, 135], or the solution based on a planar spanner graph proposed by Bruck et al. [28, 29]. In the same work, the authors have also proved that the problem of locating the nodes using local angle information is NP-Hard.

The next section describes the use of AOA information in the SOM localization algorithm. The solution targets ad-hoc networks where only a limited number of devices are equipped with a switched beam antenna, and the coverage area of these reference nodes does not necessarily include the whole network. Preliminary simulations show interesting results. In networks with sufficient connectivity, the use of a single anchor node with a switched beam antenna can effectively replace four anchors in the corners of the network.

6.3.1 The SOM Localization Scheme Using AOA Information

The SOM technique is based on a learning algorithm that applies similar weight updates to adjacent neurons. When the weights model the position of a set of wireless nodes, SOM implements a simple and elegant solution to approximate the node positions using proximity constraints (SOM-V and SOM-A) or RSS values (SOM-R). This section proposes a modification to the algorithm capable of exploiting AOA information in the training phase of the map. The basic idea is to modify the SOM learning algorithm so that, at each iteration, the position of nodes in proximity of a directional antenna is adjusted to match their estimated AOA.

System Model

Assume the typical 2D localization scenario with n nodes placed at unknown location and m anchors located at known positions. Also assume k of the base stations equipped with a switched beam DA similar to the one described in Section 6.1. These *Directional Base Stations* (DBS) are manually aligned according to a common reference system, or equipped with a magnetometer for automatic alignment with the earth's magnetic field. Let $\text{DNB}_k = \{n_1^{(k)}, \dots, n_{k_p}^{(k)}\}$ contain the ID's of the nodes in the radio range of the k^{th} DBS, and let DAOA_k contains the estimated AOA of the messages transmitted by each of such neighbor: $\text{DAOA}_k = \{\theta_1^{(k)}, \dots, \theta_{k_p}^{(k)}\}$.

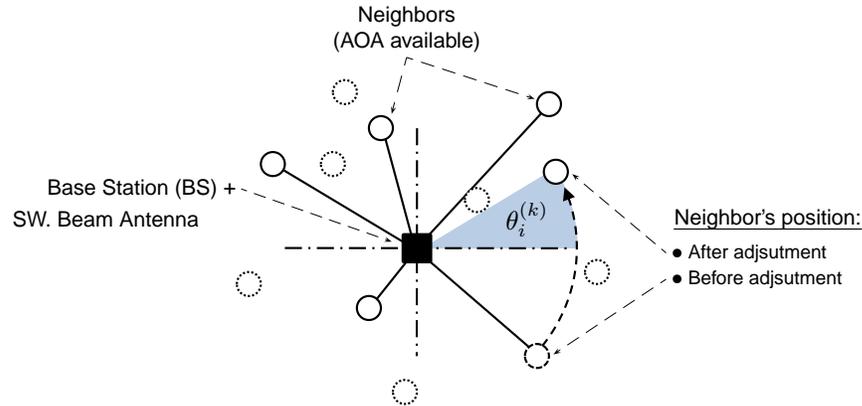


Figure 6.15: At each step of the SOM- X_{θ_k} algorithm the position of each neighbor in proximity of a directional antenna is updated to match their AOA estimate.

Modified SOM Algorithm

The base SOM algorithm is modified to include directional information in the training phase of the map. Before executing each iteration, the position of each node $n_i^{(k)}$ in the radio range of the k^{th} DBS is adjusted to match the estimated AOA $\theta_i^{(k)}$ (see Figure 6.15). Note that the adjustment only changes the bearing of each neighbor without altering their distance from the DBS. This modification is compatible both with the use of connectivity information (SOM-V and SOM-A) and RSS values (SOM-R). In the rest of this section, the notation SOM- X_{θ_k} will be used to denote a SOM variant used and the number of DBS's. For example, SOM-A $_{\theta_4}$ will indicate the variant that uses connectivity information and four DBS's, while SOM-R $_{\theta_1}$ will denote a version that use RSS values and a single DBS. The pseudocode describing the changes necessary to include AOA information in the SOM learning algorithm is shown in Algorithm 2. Extension to the 3D case is straightforward using a switched beam antenna as the one in Section 6.2 capable of estimating both the azimuth and elevation AOA of the messages.

6.3.2 Collaborative Single-Anchor Localization

This section evaluates the effectiveness of the proposed solution in a special case where the network contains only one BSD placed in the center of the network at a known location. As discussed in Section 6.2, single-anchor localization can be useful in supporting emergency response applications or any other deployments where it is not possible to install and maintain a network of reference nodes.

Anchor-free localization is evaluated first by comparing the performance of SOM-A and SOM-A $_{\theta_1}$. Note that the SOM-A $_{\theta_1}$ learning algorithm requires knowledge of the physical dimension of the deployment area. Since only one anchor is used, this information cannot be inferred from the position of anchor nodes

Algorithm 2: 2D SOM- θ Localization

Input: Matrix D_h : hop count distances among nodes

Input: Dimensions of the deployment area

Output: $[x_j, y_j]$ for $j = 1, \dots, N$: node positions

```

% Parameter Initialization
1:  $\eta_{\max} = 0.1$ ;            $\eta_{\min} = 0.01$ ;
2:  $\sigma_{\max} = \max_{i,j}\{D_h\}/2$ ;   $\sigma_{\min} = 0.001$ 

3: for all nodes  $n$  do
4:    $[x_n, y_n]^T = \text{random}()$ 
5: end for

% Main Loop
6: for  $n = 1$  : to N_ITER do
7:    $\eta = \eta_{\max} - n(\eta_{\max} - \eta_{\min}) / (N\_ITER - 1)$ 
8:    $\sigma = \sigma_{\max} - n(\sigma_{\max} - \sigma_{\min}) / (N\_ITER - 1)$ 

   % Use AOA Information to reposition DBS's neighbors
9:   for all directional base stations  $k$  do
10:    for all neighbors  $i$  of  $k$  do
11:       $dist = \|[x_i, y_i] - [x_k, y_k]\|$ 
12:       $[x_i, y_i] = [x_k + dist \cos \theta_i^{(k)}, y_k + dist \sin \theta_i^{(k)}]$ 
13:    end for
14:   end for

15:    $(x, y) = \text{random}()$  % inside the deployment area
16:    $c = \arg \min_j \|(x, y) - (x_j, y_j)\|$ 
17:   for all network nodes  $j$  do
18:      $h_{cj} = \exp(-D_h(c, j)^2 / 2\sigma^2)$ 
19:      $[x_j, y_j] += \eta h_{cj} ([x, y] - [x_j, y_j])$ 
20:   end for
21: end for

```

The gray markers show the code that has been added or modified with respect to the original version.

located in the corners of the network, as in the case of the standard SOM-A; therefore the size of the deployment area must be explicitly supplied at runtime. Also, the adjustment shown in 2 do not need to be repeated at every training cycle. The results described in the following sections were achieved by adjusting the AOA once every 25 iterations of the standard learning algorithm.

Range-Free Collaborative Localization

The first simulation set compares the performance of SOM-A and SOM-A $_{\theta 1}$ in localizing 64 node networks similar to the ones in Figure 6.16. The connectivity information was generated by binary quantization of RSS

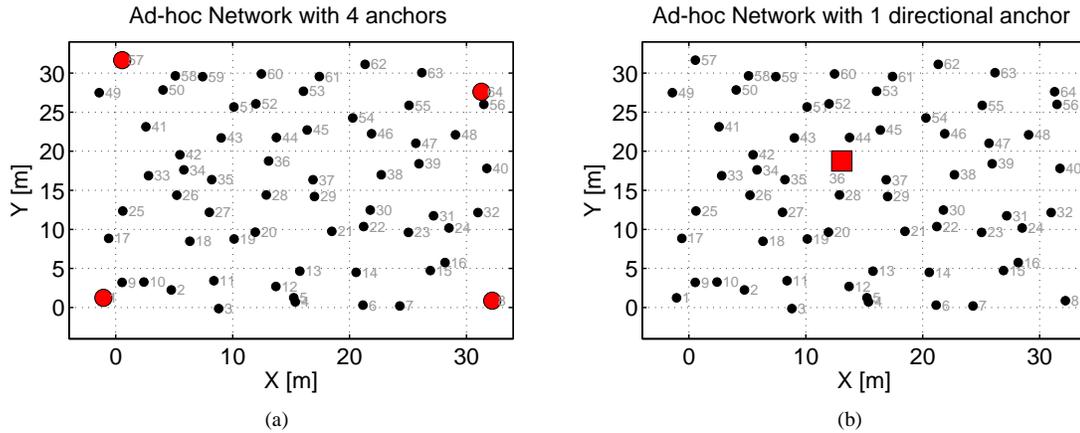


Figure 6.16: The two localization scenarios considered: a) a network with four anchors in the corners is localized using SOM-A; b) The same network is localized using SOM- A_{θ_1} ; the four anchors are replaced by a single base station capable of estimating the AOA of the messages transmitted by its neighbors.

data sampled from a log-normal distribution computed using parameters ($d_0 = 1$ m, $P_0 = -45$ dBm, $n_p = 3$, $\sigma_{dB} = 6$ dBm). Four anchor nodes were used for SOM-A, while only one anchor node placed approximately in the center of the network was used to train SOM- A_{θ_1} .

For each network localized, the SOM- A_{θ_1} 's results are evaluated using a two-step performance metric. First, the alignment of the computed map is checked to determine if it matches the ground truth. Second, the evaluation procedure computes the average localization error of the correctly aligned maps. Figure 6.17 shows the results achieved in localizing sets of 100 random topologies generated with increasing connectivity levels. For sparse networks, the number of nodes in the radio range of the DBS is low. Therefore, the angle

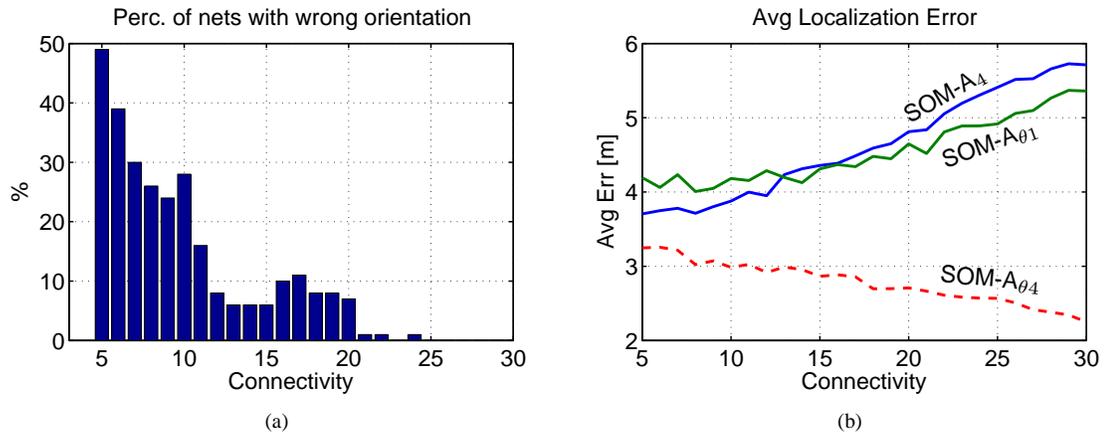


Figure 6.17: Localization results using SOM- A_4 (four anchors), SOM- A_{θ_1} (one directional anchor) and SOM- A_{θ_4} (four directional anchors).

adjustments made using the available information are not sufficient to ensure convergence to properly aligned maps. To sensibly reduce the number of incorrect topologies, the network connectivity needs to increase to values above 20. In these dense configurations, the SOM- A_{θ_1} algorithm not only produces a high percentage of aligned maps, but also achieves a lower error than SOM-A (see Figure 6.17b).

To provide a comparison baseline, the plots in Figures 6.17b also show the error achieved when the four anchor nodes in the corners are all equipped with directional antennas (SOM- A_{θ_4}). The SOM- A_{θ_4} 's error is significantly lower than the error of the other two variants and decreases steadily with the connectivity. The SOM- A_{θ_4} algorithm does not specifically take into consideration nodes that are neighbors of multiple DBSs, but when this happens, combining two AOA values is sufficient to estimate the position of the node independently from the rest of the network. When the connectivity increases, the availability of AOA estimates for a large number of nodes compensates for the loss of accuracy that affects range-free localization. As discussed earlier, several applications using directional beacons have been discussed in the literature; therefore the remaining sections will specifically focus on single-anchor applications of the directional SOM.

6.3.3 RSS-Based Collaborative Localization

As shown in the previous section, the SOM- A_{θ_1} 's performance is heavily affected by the network connectivity. A large number of neighbors in proximity of the DBS is required to ensure convergence to correctly aligned maps. Unfortunately, the need for high connectivity penalizes the accuracy of the SOM technique. As discussed in details in Chapter 4, the information available to a range-free scheme decreases in dense deployments. In fact, Figure 6.17 shows a SOM-A and SOM- A_{θ_1} 's error steadily increasing for connectivity above ten.

The limitations of range-free localization can be avoided by including RSS values in the training phase of the map (SOM-R). The simulations of the previous section have been repeated using the SOM-R scheme and its angle based variant. The number of incorrect topologies is generally lower than in the previous case, and it reduces to only a few percent for connectivity above 15 (see Figure 6.18a). The error plot shows a SOM- R_{θ_1} 's error approximately 20% lower than the SOM-R's error for the larger connectivity value tested (see Figure 6.18a).

Effect of Noisy Measurements

All the simulations presented so far were executed assuming error-free AOA estimates. To test the effect of noisy measurements, the simulations have been repeated by adding a component error err_s to each an-

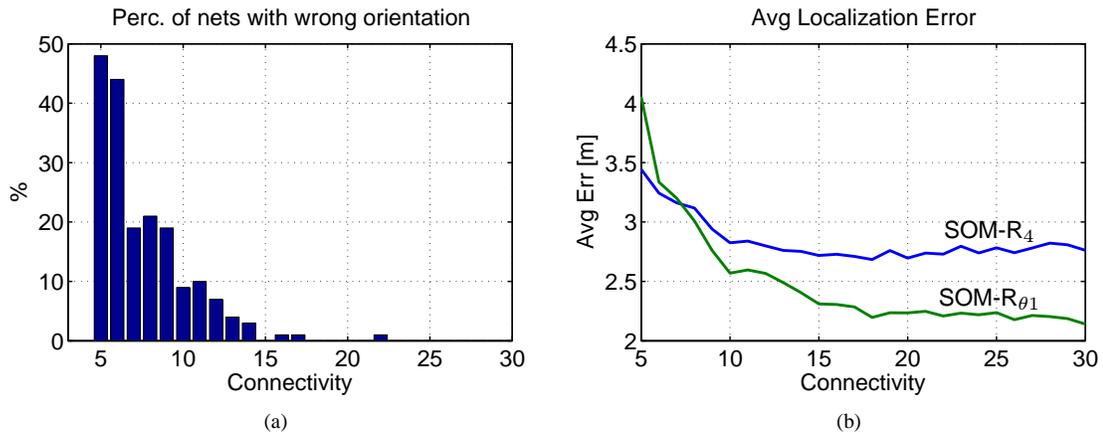


Figure 6.18: Localization results using SOM-R (four anchors) and SOM-R_{θ₁} (single anchor).

gle estimate. The error added was sampled uniformly in the interval $[-s, s]$ for four different s values: $s = \{0^\circ, 5^\circ, 10^\circ, 15^\circ\}$. These error values are compatible with the error measured during the experiment described in Section 6.1.4.

The results in Figure 6.19a do not show a significant correlation between the magnitude of the error and the number of incorrect topologies. The effect of noisy measurements is more appreciable on the error plots in Figure 6.19b, which show graceful degradation of the SOM-R_{θ₁}'s performance for increasing error values. For angular error sampled in the interval $[-15^\circ, 15^\circ]$, the SOM-R_{θ₁}'s error using a single directional anchor is close to the error of the SOM-R scheme that uses four anchors.

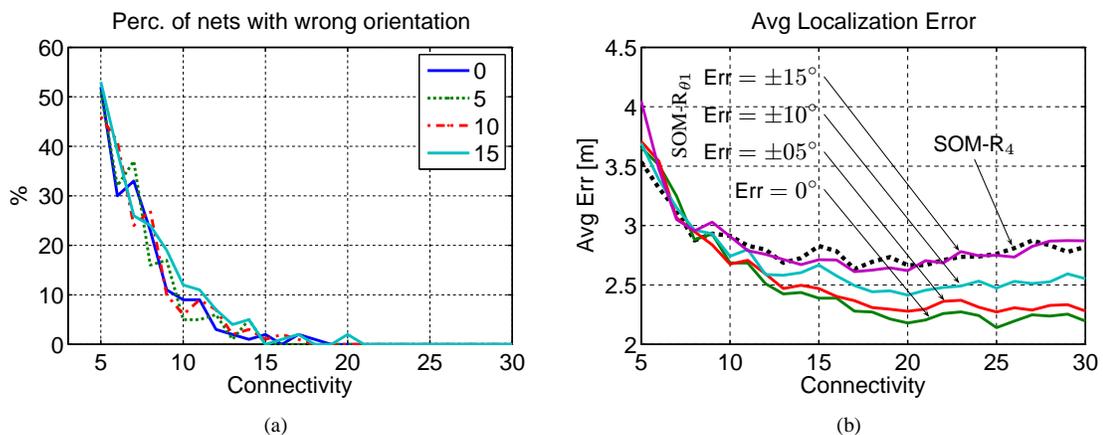


Figure 6.19: Performance of SOM-R_{θ₁} evaluated with different error affecting the AOA estimates. The black dotted line in plot (b) is the SOM-R's error with four anchors.

Chapter 7

Conclusions

Localization in ad-hoc networks requires computing node positions with only a limited amount of initial information. Chapter 3 has presented a solution that uses the Self-Organizing Map formalism to localize the nodes using radio connectivity data and (possibly) no anchor nodes. The presented solutions were able to produce accurate results in a variety of simulated scenarios. In addition, validation using RSS traces from real deployments has shown accurate localization results, especially when using the SOM-R scheme capable of combining connectivity and RSS information (see Chapter 5). This solution ensures accurate localization in sparse deployments as well as in fully connected networks; it is robust to noisy radio measurements and conditions of anisotropic layout. Finally, the SOM localization approach, although centralized, was shown to be characterized by a lightweight implementation that makes it suitable for devices with limited computational resources.

This dissertation has also addressed localization from a theoretical point of view. Sections 4.1 and 4.2 have focused on the problem of determining an optimal quantization threshold to convert the RSS values into connectivity data. As a result of an information-theoretical analysis, the optimal threshold has been shown to be related to a network connectivity value dubbed **optimal connectivity** (OC) (see Section 4.2). Notably, the OC value can be approximated using a function that only depends on the number of network nodes. Inferring connectivity on the basis of the OC value ensures a condition of maximum information content in the measurements, thus potentially reducing the error of any scheme that operates using radio proximity data.

After having defined how to convert RSS into connectivity data, Sections 4.3 and 4.4 have investigated the difference between the range-free and the range-based localization approaches. The goal was to understand in which conditions a connectivity based scheme can potentially outperform a range-based one and vice versa. Using an approach similar to the one in Sections 4.1 and 4.2, the choice between the two approaches has been shown to depend on the **critical connectivity** (CC) value described in Section 4.4. This value can be approximated using a function that depends on the network size and the ratio σ_{dB}/n_p . Similar to the OC value, knowledge of CC can reduce localization error by guiding the choice between a range-free and a range-based solution.

The theoretical results discussed in Chapter 4 are based on analysis of the CRB for the localization error. Not every range-free scheme will achieve its lowest error for network connectivity equal to OC, and RSS ranging schemes might perform worse than range-free schemes for connectivity above CC. However, if the schemes considered are known to perform close to the CRBs, analysis of the OC and CC values will provide valuable information to reduce the localization error.

Finally, the last part of this dissertation has described the use of directional antennas for AOA estimation. Section 6.2 has demonstrated the use of a *semi-dodecahedron* antenna for single-anchor localization. This novel approach takes full advantage of the qualities of AOA-based localization, and it is capable of computing the location of a 2D target without requiring knowledge of the propagation model for the RF signal. The concept of single-anchor localization has been further explored by integrating angle information in the SOM localization approach. The proposed solutions (SOM- X_{θ_1}) implement collaborative localization schemes that exploit AOA data from a single BS to produce absolute maps. Results of preliminary simulations show that in conditions of sufficient connectivity, the single-anchor version can improve over the performance of the scheme using four anchor nodes.

7.1 Future Work

Localization Using Self-Organizing Maps

Localization based on the SOM technique has been explored in detail and evaluated using extensive simulation. Where possible, data from actual sensor networks has also been used to characterize the SOM's error. The deployment of a large test bed capable of collecting real-time data would help in obtaining a better validation of the SOM's performance under realistic conditions and using different types of information (connectivity, RSS and AOA data). Another direction for future work is the design of a distributed version

of the algorithm. Although the experimental results in Section 3.7 have shown compatibility with resource-constrained devices, a distributed implementation would make SOM localization attractive to a wider range of applications.

Theoretical Analysis of RF-Based Localization Systems

The theoretical analysis in Chapter 4 is valid for schemes that use only range estimates or connectivity information. Additional information can be used with both of the approaches. For example, range-based schemes can impose constraints on the minimum separation distance between disconnected nodes (e.g. [91]). Similarly, connectivity-based schemes can use RSS information to “sort” one-hop neighbors [99]. This is also the approach used to improve the performance of the SOM-R algorithm in Section 5.2. In both the range-free and range-based cases, using additional information will cause the localization error to decrease, and the OC and CC will not necessarily provide useful information. Analysis of schemes using hybrid measurements is left for future research work.

Chapter 4 has also shown that the optimal connectivity for range-free localization can be computed using a function of the network size. This result bears similitude to the popular work of Xue and Kumar who have investigated the number of neighbors needed to ensure connectivity in a multi-hop wireless network [174]. Their results show that asymptotic connectivity is ensured when the number of neighbors per each node is $\Theta(\log n)$, where n is the network size. A study investigating further analogies between range-free localization and asymptotic connectivity is also left for future work

Another research effort can be directed toward extending the CRB analysis to more realistic radio models. The analysis in Chapter 4 assumes a log-normal shadowing model with Gaussian distribution for the average RSS. Although the use of this model is supported by both theoretical and experimental evidences, its application does not take into account the non-ideality of the transceivers used in current wireless networks. Messages transmitted using low-power devices are subject to error when the received power falls below the radio’s sensitivity. When errors occur, the messages are dropped by the hardware and carry no RSS data useful for localization. As a result, statistics based on the successfully received radio messages will be invariably biased, especially when the RSS is close to the lower limit imposed by the hardware sensitivity. Even if the underlying model was indeed log-normal, the data collected by a realistic device will not follow the same distribution. Addressing these phenomena requires to consider the rapid fluctuations in the received power (small-scale fading models [143]), compute the bit error rate, evaluate the probability that a packet is dropped, and adjust the measurement model to take into account the effect of missing data packets. Investigating the

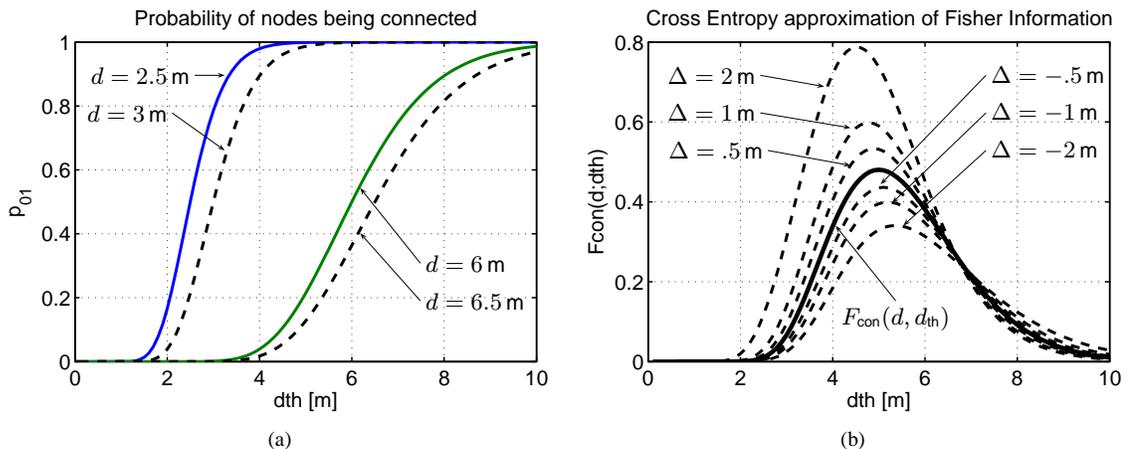


Figure 7.1: a) Probability of “connected nodes” for pairs of nodes placed at distance 2.5 m and 6 m; the same probability are also evaluated when the node positions are shifted by a factor $\Delta = 0.5$ m. b) Approximation of the Fisher information $F_{con}(d = 5 \text{ m}, d_{th})$ computed by evaluating the cross entropy $D(f \parallel f_{\Delta})$ between the probability mass functions computed with $d = 5$ m and $d = 5 + \Delta$ m, where $\Delta = \{-2, -1, -0.5, +0.5, 1, 2\}$ m.

fundamental limits under realistic operative conditions would give a significant contribution to the successful application of RF-based localization technology.

Finally, future research work could use the relation between the Fisher information and *Cross Entropy* (CE) and investigate applications to localization. The cross-entropy for two discrete random variables is described by the following equation:

$$D(p \parallel q) = \sum_{x \in \mathcal{X}} p(x) \log \frac{p(x)}{q(x)}.$$

The cross entropy, also known as *Kullback-Leibler distance*, is positive value that is often used as a distance measure between two distribution; in fact $D(p \parallel q) = 0$ if and only if $p = q$. The cross-entropy is also related to the Fisher information. If $f(X; \theta)$ is the measurement model used in the estimation process, and $f_{\Delta}(c; \theta + \Delta)$ is the same probability function evaluated for a different value of the parameter θ , then the Fisher information can be expressed as a function of the cross-entropy [50]:

$$F(\theta) = \lim_{\Delta \rightarrow 0} \left(\frac{2}{\Delta^2} D(f \parallel f_{\Delta}) \right). \quad (7.1)$$

Figure 7.1a, shows the probability of measuring two nodes as connected for different threshold val-

ues. The probabilities are computed for two nodes at distances d equal to 2.5 m and 6.0 m. The same plot also reports the probabilities when the node position is altered by adding an offset $\Delta = 0.5$ m. The differences between the two probabilities are maximized for threshold values that correspond to the true node distances, a fact that concord with the results derived in Section 4. Figure 7.1b show the correspondence between the cross-entropy computed for different values of Δ and the Fisher information for connectivity measurements when nodes are placed 5 m apart. As the offset value is reduced, i.e. $\Delta \rightarrow 0$, the value $(2/\Delta^2)D(f \parallel f_\Delta)$ becomes a close approximation of the Fisher information.

Equation (7.1) suggests that a node might be able to compute the amount of Fisher information by evaluating how the connectivity changes as the position of the nearby nodes are perturbed. Although nodes cannot be physically moved, changes in the connectivity could be evaluated by perturbing the quantization threshold P_{th} . Application of this approach could result in a distributed scheme that allows each node to determine a locally optimal threshold to estimate its own position using connectivity measurements.

REFERENCES

- [1] IEEE Std 802.15.1-2005. http://standards.ieee.org/getieee802/download/802.15.1-2005_part1.pdf.
- [2] Jennic JN5139 datasheet. <http://http://www.jennic.com/support/datasheets/>.
- [3] Jennic web site. <http://www.jennic.com>.
- [4] Online document on comp.ai.neural-nets: How many kinds of Kohonen networks exist? <http://www.faqs.org/faqs/ai-faq/neural-nets/part1/preamble.html>.
- [5] TI CC2420 datasheet. <http://www.ti.com/lit/gpn/cc2420>.
- [6] Tinyos web site. <http://www.tinyos.net>.
- [7] Akyildiz, I., Su, W., Sankarasubramaniam, Y., and Cayirci, E. Wireless sensor networks: a survey. *Computer Networks*, 38(4):393–422, 2002.
- [8] Andreas Savvides, W., Sachin Adlakha, R., and Srivastava, M. On the Error Characteristics of Multihop Node Localization in Ad-Hoc Sensor Networks. *Information Processing in Sensor Networks: Second International Workshop, Ipsn 2003, Palo Alto, Ca, Usa, April 22-23, 2003: Proceedings*, 2003.
- [9] Ansoft, . HFFS. <http://www.ansoft.com/products/hf/hfss/>.
- [10] Arora, A., Dutta, P., Bapat, S., Kulathumani, V., Zhang, H., Naik, V., Mittal, V., Cao, H., Demirbas, M., Gouda, M., and others, . A line in the sand: a wireless sensor network for target detection, classification, and tracking. *Computer Networks*, 46(5):605–634, 2004.
- [11] Ash, J. and Potter, L. Sensor network localization via received signal strength measurements with directional antennas. In *Proceedings of the 2004 Allerton Conference on Communication, Control, and Computing*, pages 1861–1870, 2004.
- [12] Aspnes, J., Eren, T., Goldenberg, D., Morse, A., Whiteley, W., Yang, Y., Anderson, B., and Belhumeur, P. A theory of network localization. *IEEE Transactions on Mobile Computing*, 5(12):1663–1678, 2006.
- [13] Aspnes, J., Goldenberg, D., and Yang, Y. On the computational complexity of sensor network localization. *Proceedings of the First International Workshop on Algorithmic Aspects of Wireless Sensor Networks*, 2004.
- [14] Bahl, P. and Padmanabhan, V. 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*, 2, 2000.
- [15] Balanis, C. A. *Antenna theory analysis and design*. J. Wiley and Sons, New York, 1997.
- [16] Bandyopadhyay, S., Roy, S., and Ueda, T. *Enhancing the Performance of Ad Hoc Wireless Networks with Smart Antennas*. CRC Press, 2006.

- [17] Bannister, K., Giorgetti, G., and Gupta, S. Wireless Sensor Networking for Hot Applications: Effects of Temperature on Signal Strength, Data Collection and Localization. *Proceedings of The Fifth Workshop on Embedded Networked Sensors (HotEmnets'08)*, 2008.
- [18] Bao, L. and Garcia-Luna-Aceves, J. Transmission scheduling in ad hoc networks with directional antennas. *Proceedings of the 8th annual international conference on Mobile computing and networking*, pages 48–58, 2002.
- [19] Basu, A., Gao, J., Mitchell, J., and Sabhnani, G. Distributed localization using noisy distance and angle information. *Proceedings of the seventh ACM international symposium on Mobile ad hoc networking and computing*, pages 262–273, 2006.
- [20] Beutel, J. *Location management in wireless sensor networks. Handbook of sensor networks: compact wireless and wired sensing systems*. CRC Press, Boca Raton, Fla., 2005.
- [21] Biswas, P., Lian, T., Wang, T., and Ye, Y. Semidefinite programming based algorithms for sensor network localization. *ACM Transactions on Sensor Networks (TOSN)*, 2(2):188–220, 2006.
- [22] Biswas, P. and Ye, Y. Semidefinite programming for ad hoc wireless sensor network localization. *Proceedings of the third international symposium on Information processing in sensor networks*, pages 46–54, 2004.
- [23] Boccuzzi, J. *Signal Processing for Wireless Communications*. McGraw-Hill Professional, 2008.
- [24] Bonabeau, E. and Henaux, F. Self-Organizing Maps for Drawing Large Graphs. *Information Processing Letters*, 67(4):177–184, 1998.
- [25] Borg, I. and Groenen, P. *Modern multidimensional scaling*. Springer New York, 1997.
- [26] Breu, H. and Kirkpatrick, D. Unit disk graph recognition is NP-hard. *Computational Geometry: Theory and Applications*, 9(1-2):3–24, 1998.
- [27] Brooks, R., Ramanathan, P., and Sayeed, A. Distributed target classification and tracking in sensor networks. *Proceedings of the IEEE*, 91(8):1163–1171, 2003.
- [28] Bruck, J., Gao, J., and Jiang, A. Localization and routing in sensor networks by local angle information. *Proceedings of the 6th ACM international symposium on Mobile ad hoc networking and computing*, pages 181–192, 2005.
- [29] Bruck, J., Gao, J., and Jiang, A. Localization and routing in sensor networks by local angle information. *Transactions on Sensor Networks (TOSN)*, 2009.
- [30] Bulusu, N., Heidemann, J., and Estrin, D. GPS-less low-cost outdoor localization for very small devices. *Personal Communications, IEEE [see also IEEE Wireless Communications]*, 7(5):28–34, 2000.
- [31] Bulusu, N., Heidemann, J., and Estrin, D. Adaptive beacon placement. *Proceedings of the 21st International Conference on Distributed Computing Systems (ICDCS-21)*, pages 489–498, 2001.
- [32] Capkun, S., Cagalj, M., and Srivastava, M. Securing localization with hidden and mobile base stations. *Proceedings of IEEE INFOCOM*, 2006.
- [33] Cayirci, E. and Cuplu, T. SENDROM: Sensor Networks for Disaster Relief Operations Management. *Proc of Med-Hoc Net 2004, the Third Annual Mediterranean Ad Hoc Networking Workshop.*, 2004.
- [34] Chang, C. and Sahai, A. Cramér-Rao-type bounds for localization. *EURASIP Journal on Applied Signal Processing*, 2006(1):166–166, 2006.

- [35] Chintalapudi, K., Dhariwal, A., Govindan, R., and Sukhatme, G. Ad-hoc localization using ranging and sectoring. *INFOCOM 2004. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*, 4, 2004.
- [36] Choudhury, R. and Vaidya, N. Ad Hoc Routing Using Directional Antennas. *Illinois Computer Systems Symposium (ICSS) May 2002, UIUC*, 2002.
- [37] Choudhury, R., Yang, X., Ramanathan, R., and Vaidya, N. Using directional antennas for medium access control in ad hoc networks. *Proceedings of ACM MOBICOM*, 2, 2002.
- [38] Chung, W., Ha, D., and Ni, D. Performance evaluation of an ultra wideband radiolocation system with directional beacon. *Radio and Wireless Conference, 2004 IEEE*, pages 59–62, 2004.
- [39] Cidronali, A., Maddio, S., Giorgetti, G., Magrini, I., Gupta, S., and Manes, G. A 2.45 GHz Smart Antenna for Location-Aware Single-Anchor Indoor Applications. *Proceedings of International Microwave Symposium (IMS)*, 2009.
- [40] Costa, J., Patwari, N., and Hero III, A. Distributed multidimensional scaling with adaptive weighting for node localization in sensor networks. *ACM Trans. on Sensor Networks*, 2005.
- [41] Cottrell, M. and Fort, J. Etude d'un algorithme d'auto-organisation. *Annales de l'Institut Poincaré*, 23:1, 1987.
- [42] Cottrell, M., Fort, J., and Pagès, G. Theoretical aspects of the SOM algorithm. *Neurocomputing*, 21(1-3):119–138, 1998.
- [43] Deshmukh, S. and Miller, J. Location-Aware Computing. *Technology*, page 1, 2003.
- [44] Dimitriou, T. and Kalis, A. Efficient delivery of information in sensor networks using smart antennas. *1st International Workshop on Algorithmic Aspects of Wireless Sensor Networks (ALGOSENSORS 2004)*, pages 109–22, 2004.
- [45] Elnahrawy, E., Austen-Francisco, J., and Martin, R. Adding angle of arrival modality to basic rss location management techniques. In *IEEE International Symposium on Wireless Pervasive Computing (ISWPCO 07), Puerto Rico*, 2007.
- [46] Eren, T., Goldenberg, D., Whiteley, W., Yang, Y., Morse, A., Anderson, B., and Belhumeur, P. Rigidity, computation, and randomization in network localization. *Proc. IEEE INFOCOM (March 2004)*, 2004.
- [47] Ertin, E. and Priddy, K. Self-localization of wireless sensor networks using self-organizing maps. *Proceedings of SPIE*, 5803:138, 2005.
- [48] Farrell, J. and Barth, M. *The Global Positioning System and Inertial Navigation*. McGraw-Hill Professional, 1999.
- [49] Fort, J. and Pages, G. On the AS Convergence of the Kohonen Algorithm with a General Neighborhood Function. *The Annals of Applied Probability*, 5(4):1177–1216, 1995.
- [50] Frieden, B. *Science from Fisher Information: A Unification*. Cambridge University Press, 2004.
- [51] Gerecke, U. and Sharkey, N. Quick and dirty localization for a lost robot. *Computational Intelligence in Robotics and Automation, 1999. CIRA'99. Proceedings. 1999 IEEE International Symposium on*, pages 262–267, 1999.
- [52] Ghosh, J., Delampady, M., and Samanta, T. *An introduction to bayesian analysis: theory and methods*. Springer Verlag, 2006.
- [53] Giorgetti, G., Cidronali, A., Gupta, K., and Manes, G. Single-anchor indoor localization using a switched-beam antenna. *Communications Letters, IEEE*, 13(1), January 2009.

- [54] Giorgetti, G., Cidronali, A., Gupta, S., and Manes, G. Exploiting Low-Cost Directional Antennas in 2.4 GHz IEEE 802.15.4 Wireless Sensor Networks. *The 37th European Microwave Conference, 2007, Munich, Germany*, 2007.
- [55] Giorgetti, G., Gupta, S., and Manes, G. Wireless localization using self-organizing maps. *Proceedings of the 6th international conference on Information processing in sensor networks*, pages 293–302, 2007.
- [56] Giorgetti, G., Gupta, S. K., and Manes, G. Localization using signal strength: to range or not to range? In *MELT '08: Proceedings of the first ACM international workshop on Mobile entity localization and tracking in GPS-less environments*, pages 91–96, New York, NY, USA, 2008. ACM.
- [57] Giorgetti, G., Gupta, S. K., and Manes, G. Optimal rss threshold selection in connectivity-based localization schemes. In *MSWiM '08: Proceedings of the 11th international symposium on Modeling, analysis and simulation of wireless and mobile systems*, pages 220–228, New York, NY, USA, 2008. ACM.
- [58] Giorgetti, G., Maddio, S., Cidronali, A., Gupta, S., and Manes, G. Switched Beam Antenna Design Principles for Angle of Arrival Estimation. *Wireless Technologies, 2009 European Conference on*, Oct. 2009.
- [59] Giorgetti, G., Manes, G., Lewis, J. H., Mastroianni, S. T., and Gupta, S. K. S. The personal sensor network: a user-centric monitoring solution. In *BodyNets '07: Proceedings of the ICST 2nd international conference on Body area networks*, pages 1–2, ICST, Brussels, Belgium, Belgium, 2007. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).
- [60] Goldenberg, D., Krishnamurthy, A., Maness, W., Yang, Y., Young, A., Morse, A., Savvides, A., and Anderson, B. Network localization in partially localizable networks. *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, 1, 2005.
- [61] Goldsmith, A. *Wireless communications*. Cambridge University Press, 2005.
- [62] Graziosi, F. and Santucci, F. A general correlation model for shadow fading in mobile radiosystems. *IEEE Communications Letters*, 6(3):102–104, 2002.
- [63] Gudmundson, M. Correlation model for shadow fading in mobile radio systems. *Electronics letters*, 27(23):2145–2146, 1991.
- [64] He, T., Huang, C., Blum, B., Stankovic, J., and Abdelzaher, T. Range-free localization schemes for large scale sensor networks. *Proceedings of the 9th annual international conference on Mobile computing and networking*, pages 81–95, 2003.
- [65] Helal, S., Mann, W., El-Zabadani, H., King, J., Kaddoura, Y., and Jansen, E. The Gator Tech Smart House: A Programmable Pervasive Space. *COMPUTER*, pages 50–60, 2005.
- [66] Hendrickson, B. Conditions for Unique Graph Realizations. *SIAM J. Comput.*, 21(1):65–84, 1992.
- [67] Hightower, J. and Borriello, G. Location systems for ubiquitous computing. *Computer*, 34(8):57–66, 2001.
- [68] Hofmann-Wellenhof, B., Lichtenegger, H., and Collins, J. *Global Positioning System: Theory and Practice*. Springer, 1993.
- [69] Huang, Z., Shen, C., Srisathapornphat, C., and Jaikaeo, C. Topology control for ad hoc networks with directional antennas. *Computer Communications and Networks, 2002. Proceedings. Eleventh International Conference on*, pages 16–21, 2002.

- [70] Intanagonwiwat, C., Govindan, R., and Estrin, D. Directed diffusion: A scalable and robust communication paradigm for sensor networks. *Proceedings of the ACM/IEEE International Conference on Mobile Computing and Networking*, pages 56–67, 2000.
- [71] Jackson, B. and Jordan, T. Connected rigidity matroids and unique realizations of graphs. *Manuscript, March*, 2003.
- [72] Jacobs, D. and Hendrickson, B. An algorithm for two-dimensional rigidity percolation: the pebble game. *Journal of Computational Physics*, 137(2):346–365, 1997.
- [73] Janet, J., Gutierrez, R., Chase, T., White, M., and Sutton, J. Autonomous mobile robot global self-localization using Kohonen and region-feature neural networks. *Journal of Robotic Systems*, 14(4):263–282, 1997.
- [74] Ji, X. and Zha, H. Sensor positioning in wireless ad-hoc sensor networks using multidimensional scaling. *INFOCOM 2004. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*, 4, 2004.
- [75] Kajiwara, A. Line-of-sight indoor radio communication using circular polarized waves. *Vehicular Technology, IEEE Trans. on*, 44(3):487–493, 1995.
- [76] Kang, I. and Poovendran, R. Design issues on broadcast routing algorithms using realistic cost-effective smart antenna models. *Vehicular Technology Conference, 2004. VTC 2004-Spring, 2004 IEEE 59th*, 4, 2004.
- [77] Kaplan, E. *Understanding GPS: Principles and Applications*. Artech House, 1996.
- [78] Karp, B. and Kung, H. GPSR: greedy perimeter stateless routing for wireless networks. *Proceedings of the 6th annual international conference on Mobile computing and networking*, pages 243–254, 2000.
- [79] Kaski, S., Kangas, J., and Kohonen, T. Bibliography of self-organizing map (SOM) papers: 1981–1997. *Neural Computing Surveys*, 1(3&4):1–176, 1998.
- [80] Kleinrock, L. and Silvester, J. Optimum transmission radii for packet radio networks or why six is a magic number. *NTC'78; National Telecommunications Conference, Birmingham, Ala., December 3-6, 1978, Conference Record.*, 1:A79–40501, 1978.
- [81] Klingenbrunn, T. and Mogensen, P. Modelling cross-correlated shadowing in network simulations. In *IEEE VTS 50th Vehicular Technology Conference, 1999. VTC 1999-Fall*, volume 3, 1999.
- [82] Kohonen, T. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43(1):59–69, 1982.
- [83] Kohonen, T. Things you haven't heard about the self-organizing map. *Neural Networks, 1993., IEEE International Conference on*, pages 1147–1156, 1993.
- [84] Kohonen, T. *Self-Organizing Maps*. Springer, 2001.
- [85] Korakis, T., Jakllari, G., and Tassiulas, L. A MAC protocol for full exploitation of directional antennas in ad-hoc wireless networks. *Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing*, pages 98–107, 2003.
- [86] Koubâa, A., Alves, M., and Tovar, E. IEEE 802.15. 4 for wireless sensor networks: a technical overview. *IPP-HURRAY Technical Report (TR-050702)*, 2005.
- [87] Krim, H. and Viberg, M. Two decades of array signal processing research: the parametric approach. *Signal Processing Magazine, IEEE*, 13(4):67–94, 1996.

- [88] Krishnamachari, B., Wicker, S., Bejar, R., and Pearlman, M. Communications, information and network security, ch. Critical Density Thresholds in Distributed Wireless Networks, 2002.
- [89] Kuhn, F., Moscibroda, T., and Wattenhofer, R. Unit disk graph approximation. *Proceedings of the 2004 joint workshop on Foundations of mobile computing*, pages 17–23, 2004.
- [90] Kuhn, F., Wattenhofer, R., Zhang, Y., and Zollinger, A. Geometric ad-hoc routing: of theory and practice. *Proceedings of the twenty-second annual symposium on Principles of distributed computing*, pages 63–72, 2003.
- [91] Kwon, Y., Mechtov, K., Sundresh, S., Kim, W., and Agha, G. Resilient Localization for Sensor Networks in Outdoor Environments. *Urbana*, 51:61801.
- [92] Laman, G. On graphs and rigidity of plane skeletal structures. *Journal of Engineering Mathematics*, 4(4):331–340, 1970.
- [93] Langendoen, K. and Reijers, N. Distributed localization in wireless sensor networks: a quantitative comparison. *Computer Networks*, 43(4):499–518, 2003.
- [94] Lazos, L. and Poovendran, R. SeRLoc: Robust localization for wireless sensor networks. *ACM Transactions on Sensor Networks (TOSN)*, 1(1):73–100, 2005.
- [95] Lazos, L., Poovendran, R., and Čapkun, S. ROPE: robust position estimation in wireless sensor networks. *Proceedings of the 4th international symposium on Information processing in sensor networks*, 2005.
- [96] Leang, D. and Kalis, A. Smart SensorDVBs: Sensor Network Development Boards with Smart Antennas. *International Conference on Communications, Circuits and Systems*, 2:1476–1480, 2004.
- [97] Lee, U., Zhou, B., Gerla, M., Magistretti, E., Bellavista, P., and Corradi, A. MobEyes: Smart Mobs for Urban Monitoring with a Vehicular Sensor Network. *IEEE WIRELESS COMMUNICATIONS*, 13(5):52, 2006.
- [98] Li, J., Jannotti, J., De Couto, D., Karger, D., and Morris, R. *A scalable location service for geographic ad hoc routing*. ACM Press New York, NY, USA, 2000.
- [99] Li, X., Shi, H., and Shang, Y. A partial-range-aware localization algorithm for ad-hoc wireless sensor networks. *Local Computer Networks, 2004. 29th Annual IEEE International Conference on*, pages 77–83, 2004.
- [100] Li, X., Shi, H., and Shang, Y. Selective anchor placement algorithm for ad-hoc wireless sensor networks. In *Communications, 2008. ICC '08. IEEE International Conference on*, pages 2359–2363, May 2008.
- [101] Li, Z., Trappe, W., Zhang, Y., and Nath, B. Robust statistical methods for securing wireless localization in sensor networks. *Proceedings of the 4th international symposium on Information processing in sensor networks*, 2005.
- [102] Liu, C., Wu, K., and He, T. Sensor localization with Ring Overlapping based on Comparison of Received Signal Strength Indicator. *Mobile Ad-hoc and Sensor Systems, 2004 IEEE International Conference on*, pages 516–518, 2004.
- [103] Liu, D., Ning, P., and Du, W. Attack-resistant location estimation in sensor networks. *Information Processing in Sensor Networks, 2005. IPSN 2005. Fourth International Symposium on*, pages 99–106, 2005.
- [104] Liu, H., Darabi, H., Banerjee, P., and Liu, J. Survey of Wireless Indoor Positioning Techniques and Systems. *Systems Man And Cybernetics, IEEE Trans. on*, 37(6):1067, 2007.

- [105] Liu, T., Sadler, C., Zhang, P., and Martonosi, M. Implementing software on resource-constrained mobile sensors: experiences with Impala and ZebraNet. In *Proceedings of the 2nd international conference on Mobile systems, applications, and services*, pages 256–269. ACM New York, NY, USA, 2004.
- [106] Lorincz, K. and Welsh, M. MoteTrack: A Robust, Decentralized Approach to RF-Based Location Tracking. *Proceedings of Intl. Workshop on Location and Context-Awareness*, 2005.
- [107] Lorincz, K. and Welsh, M. MoteTrack: a robust, decentralized approach to RF-based location tracking. *Personal and Ubiquitous Computing*, 11(6):489–503, 2007.
- [108] Lotker, Z., Albeniz, M.de , and Perennes, S. Range-Free Ranking in Sensors Networks and Its Applications to Localization. *Proceedings of 3rd Ad-Hoc, Mobile, and Wireless Networks: Third International Conference (ADHOC-NOW)*, pages 158–171, 2004.
- [109] Lymberopoulos, D., Lindsey, Q., and Savvides, A. An Empirical Analysis of Radio Signal Strength Variability in IEEE 802.15. 4 Networks using Monopole Antennas, ENALAB Technical Report 050501. *Embedded Networks and Applications Lab (ENALAB)*. Yale University New Haven, CT, 6511.
- [110] Lymberopoulos, D., Lindsey, Q., and Savvides, A. An Empirical Characterization of Radio Signal Strength Variability in 3-D IEEE 802.15. 4 Networks Using Monopole Antennas. *EWSN*, pages 326–341, 2006.
- [111] Malhotra, N., Krasniewski, M., Yang, C., Bagchi, S., and Chappell, W. Location Estimation in Ad Hoc Networks with Directional Antennas. *ICDCS*, 25:633, 2005.
- [112] Maroti, M. and Völgyesi, P. and Dóra, S. and Kusý, B. and Nadas, A. and Lédeczi, Á. and Balogh, G. and Molnar, K., . Radio interferometric geolocation. In *Proceedings of the 3rd international conference on Embedded networked sensor systems*, pages 1–12. ACM New York, NY, USA, 2005.
- [113] McErlean, D. and Narayanan, S. Distributed detection and tracking in sensor networks. In *ASILOMAR CONFERENCE ON SIGNALS SYSTEMS AND COMPUTERS*, volume 2, pages 1174–1178. IEEE; 1998, 2002.
- [114] Meyer, B. Self-organizing graphs-a neural network perspective of graph layout. *Proceedings of the 6th International Symposium on Graph Drawing*, pages 246–262, 1998.
- [115] Moore, D., Leonard, J., Rus, D., and Teller, S. Robust distributed network localization with noisy range measurements. *Proceedings of the 2nd international conference on Embedded networked sensor systems*, pages 50–61, 2004.
- [116] Moscibroda, T., O’Dell, R., Wattenhofer, M., and Wattenhofer, R. Virtual coordinates for ad hoc and sensor networks. *Proceedings of the 2004 joint workshop on Foundations of mobile computing*, pages 8–16, 2004.
- [117] Moses, R., Krishnamurthy, D., and Patterson, R. A Self-Localization Method for Wireless Sensor Networks. *EURASIP Journal on Applied Signal Processing*, 2003(4):348–358, 2003.
- [118] Mulier, F. and Cherkassky, V. Learning rate schedules for self-organizing maps. *Pattern Recognition, 1994. Vol. 2-Conference B: Computer Vision & Image Processing., Proceedings of the 12th IAPR International. Conference on*, 2, 1994.
- [119] Muthukrishnan, K., Lijding, M., and Havinga, P. Towards Smart Surroundings: Enabling Techniques and Technologies for Localization. 2005.

- [120] Nagpal, R., Shrobe, H., and Bachrach, J. Organizing a global coordinate system from local information on an ad hoc sensor network. *Proc. of Information Processing in Sensor Networks (IPSN)*, 2003.
- [121] Nasipuri, A. and Li, K. A directionality based location discovery scheme for wireless sensor networks. *Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*, pages 105–111, 2002.
- [122] Nasipuri, A. and Najjar, R. Experimental Evaluation of an Angle Based Indoor Localization System. *Proc. of 5th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, 2006.
- [123] Newsome, J., Shi, E., Song, D., and Perrig, A. The Sybil attack in sensor networks: analysis & defenses. *Information Processing in Sensor Networks, 2004. IPSN 2004. Third International Symposium on*, pages 259–268, 2004.
- [124] Nguyen, X., Jordan, M., and Sinopoli, B. A kernel-based learning approach to ad hoc sensor network localization. *ACM Transactions on Sensor Networks (TOSN)*, 1(1):134–152, 2005.
- [125] Niculescu, D. and Nath, B. Ad hoc positioning system (APS) in GLOBECOM. *IEEE, San Antonio, November*, 2001.
- [126] Niculescu, D. and Nath, B. Ad hoc positioning system (APS) using AOA. *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies. IEEE*, 3, 2003.
- [127] Niculescu, D. and Nath, B. DV Based Positioning in Ad Hoc Networks. *Telecommunication Systems*, 22(1):267–280, 2003.
- [128] Niculescu, D. and Nath, B. DV Based Positioning in Ad Hoc Networks. *Telecommunication Systems*, 22(1):267–280, 2003.
- [129] Niculescu, D. and Nath, B. Error characteristics of ad hoc positioning systems (aps). *Proceedings of the 5th ACM international symposium on Mobile ad hoc networking and computing*, pages 20–30, 2004.
- [130] Niculescu, D. and Nath, B. VOR base stations for indoor 802.11 positioning. *MobiCom*, pages 58–69, 2004.
- [131] O’Dell, R. and Wattenhofer, R. Theoretical aspects of connectivity-based multi-hop positioning. *Theoretical Computer Science*, 344(1):47–68, 2005.
- [132] Oja, M., Kaski, S., and Kohonen, T. Bibliography of self-organizing map (SOM) papers: 1998-2001 addendum. *Neural Computing Surveys*, 3(1):1–156, 2003.
- [133] Paladina, L., Paone, M., Iellamo, G., and Puliafito, A. Self Organizing Maps for Distributed Localization in Wireless Sensor Networks. *Computers and Communications, 2007. ISCC 2007. IEEE Symposium on*, pages 1113–1118, 2007.
- [134] Patwari, N. and Agrawal, P. Effects of correlated shadowing: Connectivity, localization, and rf tomography. In *Information Processing in Sensor Networks, 2008. IPSN’08. International Conference on*, pages 82–93, 2008.
- [135] Patwari, N., Ash, J., Kyperountas, S., Hero III, A., Moses, R., and Correal, N. Locating the nodes: cooperative localization in wireless sensor networks. *Signal Processing Magazine, IEEE*, 22(4):54–69, 2005.
- [136] Patwari, N., Hero, A., Perkins, M., Correal, N., and O’Dea, R. Relative location estimation in wireless sensor networks. *IEEE Transactions on Signal Processing*, 51(8):2137–2148, 2003.

- [137] Patwari, N. and Hero III, A. Using proximity and quantized RSS for sensor localization in wireless networks. *Proceedings of the 2nd ACM international conference on Wireless sensor networks and applications*, pages 20–29, 2003.
- [138] Polastre, J., Szewczyk, R., and Culler, D. Telos: enabling ultra-low power wireless research. *Proceedings of the 4th international symposium on Information processing in sensor networks*, 2005.
- [139] Priyantha, N., Balakrishnan, H., Demaine, E., and Teller, S. Anchor-free distributed localization in sensor networks. *Proc. 1st Inter. Conf. on Embedded Networked Sensor Systems (SenSys 2003)*, pages 340–341, 2003.
- [140] Priyantha, N., Chakraborty, A., and Balakrishnan, H. The Cricket location-support system. *Proceedings of the 6th annual international conference on Mobile computing and networking*, pages 32–43, 2000.
- [141] Raghavendra, C., Sivalingam, K., and Znati, T. *Wireless sensor networks*. 2004.
- [142] Rao, A., Papadimitriou, C., Shenker, S., and Stoica, I. Geographic routing without location information. *Proceedings of the 9th annual international conference on Mobile computing and networking*, pages 96–108, 2003.
- [143] Rappaport, T. *Wireless Communications: Principles and Practice*. IEEE Press Piscataway, NJ, USA, 1996.
- [144] Romer, K. The Lighthouse Location System for Smart Dust. *MobiSys 2003*, 2003.
- [145] Rudafshani, M. and Datta, S. Localization in wireless sensor networks. In *Proceedings of the 6th international conference on Information processing in sensor networks*, pages 51–60. ACM New York, NY, USA, 2007.
- [146] Sakurai, A., Nakamura, M., Furubo, S., and Ban, H. Sensor Localization Using Self-Organizing Map for Human Tracking. *1st International Conference on Sensing Technology*, pages 311–315, 2005.
- [147] Santivanez, C. and Redi, J. On the use of directional antennas for sensor networks. *Military Communications Conference, 2003. MILCOM 2003. IEEE*, 1, 2003.
- [148] Savarese, C., Rabaey, J., and Langendoen, K. Robust positioning algorithms for distributed ad-hoc wireless sensor networks. *USENIX Technical Annual Conference*, pages 317–328, 2002.
- [149] Savvides, A., Garber, W., Adlakha, S., Moses, R., and Srivastava, M. On the error characteristics of multihop node localization in ad-hoc sensor networks. *Proc. IPSN (Palo Alto, CA, April 2003)*, pages 317–332, 2003.
- [150] Savvides, A., Han, C., and Srivastava, M. Dynamic fine-grained localization in Ad-Hoc networks of sensors. *Proceedings of the 7th annual international conference on Mobile computing and networking*, pages 166–179, 2001.
- [151] Savvides, A., Park, H., and Srivastava, M. The bits and flops of the n-hop multilateration primitive for node localization problems. *Proceedings of the First ACM International Workshop on Wireless Sensor Networks and Applications (WSNA), Atlanta, GA*, pages 112–121, 2002.
- [152] Saxe, J. Embeddability of weighted graphs in k-space is strongly NP-hard. *Proc. 17th Allerton Conf. Commun. Control Comput.*, pages 480–489, 1979.
- [153] Shang, Y. and Ruml, W. Improved MDS-based localization. *INFOCOM 2004. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*, 4, 2004.

- [154] Shang, Y., Ruml, W., Zhang, Y., and Fromherz, M. Localization from mere connectivity. *Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing*, pages 201–212, 2003.
- [155] Shang, Y., Shi, H., and Ahmed, A. Performance study of localization methods for ad-hoc sensor networks. In *Mobile Ad-hoc and Sensor Systems, 2004 IEEE International Conference on*, pages 184–193, Oct. 2004.
- [156] Spyropoulos, A. and Raghavendra, C. Energy efficient communications in ad hoc networks using directional antennas. *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, 1, 2002.
- [157] Stafford, J., editor. *Precision Agriculture - an International Journal on Advances in the Science of Precision Agriculture*. Springer, 1999-2006.
- [158] Stoleru, R., He, T., Stankovic, J., and Luebke, D. A high-accuracy, low-cost localization system for wireless sensor networks. *Proceedings of the 3rd international conference on Embedded networked sensor systems*, pages 13–26, 2005.
- [159] Stoleru, R., He, T., and Stankovic, J. A. "Secure Localization and Time Synchronization for Wireless Sensor and Ad Hoc Networks", volume 30, chapter Range-free Localization. Springer, 2006.
- [160] Stoleru, R., He, T., and Stankovic, J. A. "Range-free Localization", chapter "Secure Localization and Time Synchronization for Wireless Sensor and Ad Hoc Networks" editors: Radha Poovendran, Cliff Wang, and Sumit Roy. *Advances in Information Security series, Vol. 30*. 2007.
- [161] Stoleru, R. and Stankovic, J. Probability grid: a location estimation scheme for wireless sensor networks. *Sensor and Ad Hoc Communications and Networks, 2004. IEEE SECON 2004. 2004 First Annual IEEE Communications Society Conference on*, pages 430–438, 2004.
- [162] Takai, M., Bagrodia, R., and Ren, A. Directional virtual carrier sensing for directional antennas in mobile ad hoc networks. *Proceedings of the 3rd ACM international symposium on Mobile ad hoc networking & computing*, pages 183–193, 2002.
- [163] Takizawa, Y., Davis, P., Kawai, M., Iwai, H., Yamaguchi, A., and Obana, S. Self-Organizing Location Estimation Method using Ad-hoc Networks. *Proceedings of the 7th International Conference on Mobile Data Management (MDM'06)-Volume 00*, 2006.
- [164] Tenenbaum, J., Silva, V., and Langford, J. A Global Geometric Framework for Nonlinear Dimensionality Reduction, 2000.
- [165] Tseng, Y., Kuo, S., Lee, H., and Huang, C. Location Tracking in a Wireless Sensor Network by Mobile Agents and Its Data Fusion Strategies. *The Computer Journal*, 47(4):448–460, 2004.
- [166] Heijden, F.van der . *Classification, Parameter Estimation and State Estimation: An Engineering Approach Using MATLAB*. Wiley, 2004.
- [167] Vivekanandan, V. and Wong, V. Ordinal MDS-based localisation for wireless sensor networks. *International Journal of Sensor Networks*, 1(3):169–178, 2006.
- [168] Wang, H., Yip, L., Yao, K., and Estrin, D. Lower bounds of localization uncertainty in sensor networks. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2004)*, 2004.
- [169] Wang, K., Proakis, J., and Rao, R. Energy-Efficient Routing Algorithms Using Directional Antennas for Mobile Ad Hoc Networks. *International Journal of Wireless Information Networks*, 9(2):105–118, 2002.

- [170] Ward, A., Jones, A., and Hopper, A. A new location technique for the active office. *Personal Communications, IEEE [see also IEEE Wireless Communications]*, 4(5):42–47, 1997.
- [171] Whitehouse, K. The Design of Calamari: an Ad-hoc Localization System for Sensor Networks. *University of California at Berkeley*, 2002.
- [172] Whitehouse, K., Karlof, C., Woo, A., Jiang, F., and Culler, D. The effects of ranging noise on multihop localization: an empirical study. *Information Processing in Sensor Networks, 2005. IPSN 2005. Fourth International Symposium on*, pages 73–80, 2005.
- [173] Xu, J., Shen, X., Mark, J., and Cai, J. Mobile location estimation for DS-CDMA systems using self-organizing maps. *Wireless communications and mobile computing(Print)*, 7(3):285–298, 2007.
- [174] Xue, F. and Kumar, P. The number of neighbors needed for connectivity of wireless networks. *Wireless Networks*, 10(2):169–181, 2004.
- [175] Yang, C., Bagchi, S., and Chappell, W. Location Tracking with Directional Antennas in Wireless Sensor Networks. *Microwave Symposium Digest, 2005 IEEE MTT-S International*, pages 131–134, 2005.
- [176] Yang, C., Mastarone, J., and Chappell, W. Directional antennas for angular diversity in wireless sensor networks. *Antennas and Propagation Society International Symposium, 2005 IEEE*, 4, 2005.
- [177] Yi, S., Pei, Y., and Kalyanaraman, S. On the capacity improvement of ad hoc wireless networks using directional antennas. *Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing*, pages 108–116, 2003.
- [178] Zâruba, G., Huber, M., Kamangar, F., and Chlamtac, I. Indoor location tracking using RSSI readings from a single Wi-Fi access point. *Wireless Networks*, 13(2):221–235, 2007.
- [179] Zhao, F. and Guibas, L. *Wireless Sensor Networks: An Information Processing Approach*. Morgan Kaufmann, 2004.