

Ontwerp van lokalisatie-algoritmen voor draadloze sensornetwerken
op basis van lineaire regressie

Design of Linear Regression Based Localization Algorithms
for Wireless Sensor Networks

Frank Vanheel

Promotoren: prof. dr. ir. I. Moerman, prof. dr. ir. J. Verhaever, prof. dr. ir. E. Laermans
Proefschrift ingediend tot het behalen van de graad van
Doctor in de Ingenieurswetenschappen

Vakgroep Informatietechnologie
Voorzitter: prof. dr. ir. D. De Zutter
Faculteit Ingenieurswetenschappen en Architectuur
Academiejaar 2012 - 2013



ISBN 978-90-8578-624-5
NUR 975, 986
Wettelijk depot: D/2013/10.500/57



Universiteit Gent
Faculteit Ingenieurswetenschappen en Architectuur
Vakgroep Informatietechnologie

Ontwerp van lokalisatie-algoritmen voor draadloze sensornetwerken op basis van lineaire regressie

Design of Linear Regression based Localization Algorithms
for Wireless Sensor Networks

Frank Vanheel



Proefschrift tot het bekomen van de graad van
Doctor in de Ingenieurswetenschappen
Academiejaar 2012-2013



Universiteit Gent
Faculteit Ingenieurswetenschappen en Architectuur
Vakgroep Informatietechnologie

Promotoren: prof. dr. ir. Ingrid Moerman
prof. dr. ir. Jo Verhaevert
prof. dr. ir. Eric Laermans

Universiteit Gent
Faculteit Ingenieurswetenschappen en Architectuur
Vakgroep Informatietechnologie
Gaston Crommenlaan 8 bus 201, B-9050 Gent, België
Tel.: +32-9-331.49.00
Fax.: +32-9-331.48.99



Proefschrift tot het behalen van de graad van
Doctor in de Ingenieurswetenschappen
Academiejaar 2012-2013

Dankwoord

Dit dankwoord komt nog voor de samenvatting van dit boek en neemt dus een voorname plaats in, al was het maar omdat dit voor de vele potentiële lezers de eerste kennismaking met mijn werk is.

Wat je nu vasthoudt is dus mijn doctoraatsthesis. Dit is de kroon op het werk van zeven jaar intense studie en inzet. In mijn achtenvijftigste levensjaar ben ik nu zelf (doctoraats)student af. Het was een boeiende periode, waarbij uitdagingen hand in hand gingen met nieuwe inzichten.

Een doctoraat schrijven kan je niet alleen, ik maak hier dus plaats om iedereen te bedanken die bijgedragen heeft om dit doel te bereiken.

Vele technisch boeken beginnen met “To my wife”, “To my parents”, “To my daughter”, Dit werk draag ik dus gezwind op aan mijn echtgenote Mieke. Bedankt, Mieke dat je me onvoorwaardelijk steunde en als sterke vrouw steeds achter mij bleef staan!

Een heel prominente plaats nemen mijn promotoren in. De deur van mijn hoofdpromotor prof. dr. ir. Ingrid Moerman staat altijd open. Een organisatorisch probleem, praktische vragen over de aanpak van je doctoraat, goede raad omtrent de ingeslagen weg, budget voor een conferentie, een interessante discussie, waardevolle feedback, ... ik vroeg het maar en in een knip was alles zonneklaar. Bedankt, Ingrid!

Prof. dr. ir. Jo Verhaevert is mijn promotor aan de Hogeschool Gent. Deze mentor “pur sang”loodste me zonder ik het eigenlijk tenvolle besepte binnen in de onderzoekswereld! Bedankt, Jo voor het dagdagelijkse “over de schouder kijken”, de positieve feedback en de interessant technische discussies. Oprecht bedankt voor je steun en het vertrouwen!

Alle goede dingen bestaan uit drie. Mijn derde promotor, prof. dr. ir. Eric Laermans bewonder ik om zijn enorm veelzijdige technische kennis. Bedankt, Eric voor de overvloedige opbouwende kritiek bij het nalezen van papers, jouw kritisch meedenken en de talrijke boeiende discussies!

In mijn universitair biotoop vind je de bureaugenoten Bart, Eli, Evy, Jono, Lieven, Peter DV, Peter R., Pieter B., Pieter DM, Stefan en Vincent. Bedankt, collega’s niet alleen voor de vlotte babbel, maar ook voor het beantwoorden van mijn vragen over het testbed. Bij jou Bart, de geestelijke vader van dit kleinoord kon ik steeds terecht voor het inwilligen van moeilijkere verzoeken. Wel bedankt voor jouw onvermoeibare inspanningen.

Ben, David, Dimitri, Dirk, Francis, Gianni, Jan, Jo, Kris, Luc, Maarten, Patrick, Paul R., Paul D., Peter, Stefaan mogen in deze dankbetuiging zeker niet

ontbreken, gewoon omdat jullie toffe Hogent collega's zijn. Teambuilding bracht ons dichterbij elkaar. Bedankt voor de goede sfeer, waarin onderzoek echt kan gedijen. Speciale dank gaat dus ook uit naar onze vakgroepvoorzitter prof. ir. Jan Beyens. Bedankt, Jan!

Pas als je aan het schrijven van dit boek begint, besef je dat zoveel anderen je voorgegaan zijn. Bedankt, postdocs Eli, Patrick, Stefan en Wim voor de praktische tips bij het afronden van dit doctoraat.

Och ja, we kennen Murphy allemaal, een PC die het begeeft net voor je deadline, de poster die niet afgedrukt raakt, een virus op je computer, Bedankt, Dries P, Bert, Davinia, Joeri, Jonathan en Martine voor de administratieve ondersteuning.

Uiteraard mag ik prof. dr. Marc Vanhaelst, decaan van onze GTI faculteit niet vergeten. Bedankt, Marc voor de steun aan mijn doctoraatsonderzoek en de financiering ervan!

Prof. dr. ir. Piet Demeester is de drijvende kracht achter de explosief groeiende IBCN onderzoeksgroep. Hartelijk bedankt, Piet dat ik binnen jouw onderzoeksgroep boeiend onderzoek kon verrichten in een aangename en bijzonder goed uitgeruste omgeving!

Gent, september 2013
Frank Vanheel

,

Table of Contents

Dankwoord	i
Nederlandstalige samenvatting	xxiii
English Summary	xxvii
1 Introduction	1
1.1 Importance of indoor localization	1
1.2 State of the art	2
1.2.1 State-of-the-art ranging techniques	3
1.2.2 State-of-the-art location estimators	5
1.2.3 State-of-the-art wireless technologies	7
1.2.4 Statistical tools	13
1.3 Design goals for localization algorithms	13
1.4 Why Wireless Sensor Networks (WSN)	15
1.5 Outline and research contributions	17
1.6 Scope of this thesis	18
1.7 Publications	19
1.7.1 A1: Publications in International Journals	19
1.7.2 P1: Publications indexed by the ISI Web of Science “Conference Proceedings Citation Index”	19
1.7.3 C1: Articles in other conference proceedings	20
1.7.4 C3: Abstracts in conference proceedings	20
References	21
2 Physical Layer	31
2.1 Introduction	31
2.2 The test bed: WSN and RadioPerf	31
2.3 Challenges with indoor RSSI	34
2.3.1 Slow fading versus Fast fading	34
2.3.2 Multipath fading	36
2.3.3 Interference	38
2.3.3.1 Study of interference in the IEEE 802.15.4 standard	38

2.3.3.2	Frequency spectra of an IEEE 802.11 3-Com wireless Wi-Fi NIC	45
2.3.3.3	Power measurement of Wi-Fi into IEEE 802.15.4 channels	46
2.3.3.4	Minimum distance Wi-Fi transmitter and IEEE 802.15.4 receiver	47
2.3.3.5	Interference from microwave oven	54
2.3.4	Radiation pattern	55
2.4	Design of a WSN transmitter	57
2.4.1	Introduction	57
2.4.2	Hardware	58
2.4.3	Software	61
2.4.4	Results	63
2.4.5	Conclusion	64
2.5	Conclusions	66
	References	68
3	Localization	75
3.1	Introduction	75
3.2	Range-based localization steps	76
3.2.1	Conversion of RSSI to a distance	76
3.2.2	Conversion of distances to a position	77
3.3	Multilateration	77
3.3.1	Geometric multilateration	77
3.3.1.1	The geometric multilateration principle	77
3.3.1.2	Geometric trilateration with obstructed path outliers	80
3.3.1.3	Geometric trilateration with constructive multi-path outliers	81
3.3.1.4	A linearized implementation of geometric trilateration	81
3.3.2	Statistical multilateration	83
3.3.2.1	The statistical multilateration principle	83
3.3.2.2	A linearized implementation of statistical trilateration	85
3.4	Maximum likelihood	86
3.4.1	The maximum likelihood principle	86
3.4.2	False extrema and false convergences	88
3.5	Min-Max localization algorithm	88
3.5.1	The Min-Max localization principle	88
3.5.2	Min-Max positioning algorithm with obstructed path outliers	91
3.5.3	Min-Max positioning algorithm with constructive multi-path fading	93
3.6	Case study: Effect of outliers on simple localization algorithms with three anchors	93

3.7	Conclusions	96
	References	98
4	2D-localization	101
4.1	Selection of anchors	102
4.1.1	Selection of anchors based on the linearity of their calibrated path loss model	103
4.1.2	Selection of anchors with low standard error	107
4.1.3	Form factor of the building	111
4.1.4	Complexity of the calibration process and robustness against environmental changes	112
4.2	Preprocessing	113
4.2.1	Min-max algorithm	114
4.2.2	Elimination of circles	118
4.2.3	Grouping Anchors	123
4.2.4	Preprocessing results: the distance error	125
4.3	LiReFLoA	127
4.3.1	Calculate intersection points of the longitudinal borderlines of the building and the large circles	129
4.3.2	Is the amount of constructive multipath fading acceptable?	130
4.3.3	Calculate typical circles intersection points within the building	131
4.3.4	Is there more than one intersection point within the building?	131
4.3.5	Calculate initial position as the centroid of all typical circles intersection points	132
4.3.6	Calculate the initial position as the intersection of the diagonals of the trapeze	132
4.3.7	Calculate the estimated position of the nearest intersection points of the longitudinal line and the typical circles	132
4.3.8	Calculate the position as the intersection of the diagonals of the trapeze	134
4.3.9	LiReFLoA results: the positioning error	134
4.4	Synergy of the preprocessing and LiReFLoA algorithms	137
4.5	LiReCoFuL	137
4.5.1	The LiReCoFuL cost function	138
4.5.2	Theoretical comparison LiReCoFuL and RLE	140
4.5.3	Test conditions for grid based maximum likelihood	141
4.5.4	Graphical comparison of the cost functions	142
4.5.5	Cumulative distribution function plots of the position error	143
4.5.6	Execution times	144
4.5.7	The cost function with different scenarios	145
4.6	2DLiReFLoA and LiReCoFuL	146
4.6.1	2DLiReFLoA	146
4.6.2	Comparison 2DLiReFLoA and LiReCoFuL	146
4.7	Conclusions	146

References	147
5 Pseudo-3D localization	151
5.1 Introduction	151
5.2 Test conditions for pseudo-3D test	154
5.3 P3DLiReFLoA	156
5.4 Results	160
5.4.1 Comparison of the two-dimensional and pseudo-3D cali- bration	160
5.4.2 Comparison of our algorithm with a more conventional al- gorithm.	163
5.4.3 Comparison of the two-dimensional and pseudo-3D algo- rithms	166
5.5 Conclusions	167
References	167
6 Overall conclusions	171
6.1 Overall conclusions	171
6.2 Outlook and future work	173
References	175
.	178

List of Figures

2.1	Position of sensor nodes on the third floor (a) and second floor (b) of the iMinds office building. Drywall walls are presented by dark green solid lines. The solid gray lines are concrete wall. The black squares and black circles are the selected anchors (see section 4).	32
2.2	RadioPerf with its configuration tabs, selection region and display area.	34
2.3	Cumulative distribution function plots for the error on the average RSSI.	35
2.4	Capricious line of sight (LOS) RSSI multipath fading in corridors for node 2 at different power levels.	37
2.5	USA minimal interference systems consist of Wi-Fi channels 1, 6 and 11, combined with IEEE 802.15.4 channels 15, 20, 25 and 26.	41
2.6	European minimal interference systems consist of Wi-Fi channels 1, 7 and 13, combined with IEEE 802.15.4 channels 15, 16, 21 and 22.	42
2.7	Test setup for Ubiwave UW-CM-06 spectrum and power control measurements	42
2.8	IEEE 802.15.4 adjacent channels. Channel spacing is 5 MHz. For these devices, the output level is too high, due to the presence of an internal amplifier.	43
2.9	IEEE 802.15.4 DSSS power control. When saturating the CC2420 radio chip, the passband widens and some frequency dips disappear.	44
2.10	The Wi-Fi transmitted power as function of the frequency. All measured spectra fulfill its respective frequency masks.	46
2.11	Test setup with shielded QOSMOTEC boxes for testing interference of a Wi-Fi interferer on an IEEE 802.15.4 link.	50
2.12	Equivalent test setup for determining the minimum required distance separation.	51
2.13	Equivalent distance measurement setup.	52
2.14	Worst case distance required between LAN interferer and IEEE 802.15.4 receiver for frequency offsets of 3, 18 and 48 MHz.	53
2.15	Worst case distance required between LAN interferer and IEEE 802.15.4 receiver for frequency offsets of 8, 13 and 68 MHz.	53

2.16	The spectrum of a microwave oven almost spans the whole ISM 2.4 GHz frequency band and reaches maximum output around 2.46 GHz.	55
2.17	A microwave oven with a conventional power supply does not completely block an IEEE 802.15.4 link, because the microwave only causes interference on the positive peaks of the mains supply.	56
2.18	Principle of a F-shape antenna: it consists of one open and one stubbed transmission line. The parasitic capacitances to the ground plane are also shown.	57
2.19	Radiation pattern of the T-mote Sky module (top view) with horizontal mounting [45].	58
2.20	The block scheme of the hardware used.	59
2.21	The DAQmx card contains both analog and digital inputs and outputs.	59
2.22	The transmitting PCB contains the up-converter, and can easily be debugged.	60
2.23	The block scheme of the classical digital PLL.	60
2.24	The electronic design of the PLL including frequency synthesizer, loop filter and VCO.	61
2.25	The block scheme for an IEEE 802.15.4 signal.	62
2.26	Schematic view of the acknowledge frame.	62
2.27	I-signal and Q-signal as function of the time.	64
2.28	Spectrum of the I-signal (a) and Q-signal (b). These signals are input for the AD8349 Quadrature modulator.	65
2.29	The spectral output of the generated IEEE 802.15.4 signal fulfills the absolute and relative PSD requirements.	66
2.30	A Chipcon packet sniffer detects the generated ACK-signal.	67
3.1	Geometric trilateration where all distance circles are correct. The distance circles intersect in one single point: the estimated position, represented by the small red circle.	78
3.2	Illustration of the geometric trilateration principle. In (a) a global view is given, the detailed view is shown in (b). The estimated position is the centroid of the triangle with cornerpoints A, B and C. These points are the intersections of the distance circles.	79
3.3	With geometric multilateration the required number of trilaterations increases cubically with the number of anchors.	80
3.4	The geometric trilateration with an obstructed path outlier. The triangle degrades to the single line AD. The result of the linearized implementation of the trilateration (red hexagram), see section 3.3.2.2, is not correct: it falls outside the common intersection polygon with vertices A and D.	81
3.5	The geometric trilateration with a constructive multipath outlier. There is no common intersection of all the distance disks.	82

3.6	The contribution of each anchor in the (joint density) MMSE cost function	87
3.7	The MMSE cost function for the case where all distances are estimated correctly	87
3.8	A cost function in a real-life environment with both constructive multipath and obstructed path outliers: the local minimum is closer to the target than the absolute minimum [30]	89
3.9	Principle of the Min-Max localization algorithm.	90
3.10	The Min-Max localization algorithm with an obstructed path outlier.	92
3.11	The Min-Max localization algorithm with a constructive multipath outlier.	94
3.12	Comparison of different localization algorithms with three anchors for (a) correct estimated distances, (b) an obstructed path environment, (c) a constructive multipath environment, (d) an outlier obstructed path environment, and (e) an outlier constructive multipath environment.	95
4.1	Scatter plot of all reported averaged RSSI-readings in function of the distance (on a logarithmic scale)	102
4.2	RSQ correlation coefficient of RSSI versus logarithmic distance for the sending nodes on the 3rd floor. Some nodes are highly RSSI-logarithmic distance correlated.	104
4.3	RSSI versus logarithmic distance calibration of the 10 best anchors.	105
4.4	RSSI versus logarithmic distance for well behaving sensor node 43. The node has acceptable multipath fading at all receiving nodes.	105
4.5	RSSI versus logarithmic distance for bad behaving node 21. The slope is less steep.	106
4.6	Frequency distribution for different points lying on the logarithmic distance versus RSSI regression line.	108
4.7	<i>Error_on_Distance</i> (defined as two times the estimated standard deviation of the logarithmic distance frequency distribution) for the sending nodes on the 3rd floor. Some central nodes have low maximum <i>Error_on_Distance</i>	109
4.8	RSQ correlation coefficient versus <i>Error_on_Distance</i> showing a good linear fit.	109
4.9	Regression line with probability intervals and expectation intervals (also hyperbola) for anchor 30.	110
4.10	Accuracy of the regression model for two anchors placed at the extremities of the (rectangle shaped) building. The darkest intensities correspond to the greatest accuracy. Placing anchors in the center will improve accuracy.	111
4.11	Effect of the shape of a building on the longitudinal resolution. (a) rectangular building and (b) square building.	112

4.12	The blue lines represent transmission paths between node 4 and anchor 31, node 4 and anchor 47, and also node 4 and 33. These paths are not representative for the majority of the transmission paths. They have high attenuation, due to the passing through walls and the occurrence of oblique incidence on walls.	119
4.13	Drawing of (dot marked) typical circles for positioning (squared) target 4. The red circles are ignored because they are larger than the limit for the hardware.	120
4.14	Drawing of (dot marked) typical circles for positioning (squared) target 55. All measurements from anchors at the right-hand side and the center of the building have low accuracy for targets at the left-hand side. The largest (yellow) circle at the left-hand side is ignored because it contains too many other typical circles.	121
4.15	Drawing of (dot marked) typical circles for positioning (squared) target 54. The blue circles is deleted because it is contained by more than 5 other circles. The green circles are kept and can be used in the maximum likelihood algorithm on the distance.	122
4.16	Application of the maximum likelihood algorithm on the distance on the magenta group brings the (black dotted) typical circles closer to the target.	124
4.17	Log-normal probability functions for node 14. The gray solid line represents the gradient of the maximum likelihood function. The geometric average (53.4 m) of the abscissa where the respective probability functions peak is a good approximation of the abscissa of the peak of the maximum likelihood function (54.5 m).	125
4.18	Cumulative distribution function plots for the preprocessing steps. Each subsequent preprocessing step further reduces the median of the distance error.	126
4.19	Flowchart of the LiReFLoA positioning algorithm.	128
4.20	Common intersection of the large distance circles and the longitudinal borderlines of the building. Initial position is calculated as the intersection of the diagonals of the trapeze (when initial position cannot be calculated from typical circles).	129
4.21	Large distance circles in case of constructive multipath fading. Too small large circles result in trapezes with negative sides. Initial position is calculated as the intersection of the diagonals of the trapeze.	130
4.22	Initial position based on averaging of typical circles intersection points.	131
4.23	Large distance circles in case of constructive multipath fading. Too small large circles result in trapezes with negative sides. Initial position is calculated as the intersection of the diagonals of the trapeze.	133

4.24	Final step of the localization algorithm. A weighted averaging of typical circle points in the vicinity of the initial position results in the final position.	133
4.25	Cumulative distribution function plots for the error on the position. Comparison is made between our algorithm and conventional maximum likelihood methods, presented by the mean square cost function (MMSE).	135
4.26	Error on the position along the longitudinal coordinate of the building.	136
4.27	Cumulative distribution plot of the lateral error on the position shows our algorithm performs better with a median around 2.6 m (from 3.49).	137
4.28	Linear regression and distance probability distribution.	139
4.29	Position of the active nodes, the anchors and a central target for the LiReCoFuL test on the second floor.	142
4.30	Comparison of the different cost functions for a central target (a) MMSE, (b) RLE, (c) RBRLE and (d) LiReCoFuL.	143
4.31	Cumulative distribution plot of the position error for the different cost functions on the second floor.	144
4.32	Cumulative distribution function plot of the position errors for the different cost functions on the third floor.	145
4.33	Cumulative distribution function plot of the position errors for the LiReFLoA with preprocessing and the LiReCoFuL algorithms on the third floor.	146
5.1	Position of sensor nodes on the third floor of the iMinds office building. Drywall walls are presented by blue solid lines. The solid brown lines are concrete wall. The black squares are the selected anchors (see section 4.1).	155
5.2	Schematic overview of the pseudo-3D environment. All but one pseudo-3D distance (represented by solid red lines are approximated by their projection in the plane of the fixed nodes (represented by solid blue lines).	155
5.3	This flowchart illustrates the similarities and the differences between the two-dimensional and three-dimensional algorithm. . . .	157
5.4	Comparison of the correlation coefficient of the two-dimensional and pseudo-3D tests. Two-dimensional nodes with high RSQ are also pseudo-3D nodes with high RSQ.	158
5.5	Pseudo-3D regression lines not only have higher attenuation at low distances, but are also less steep than two-dimensional counterparts. The pseudo-3D calibration procedure has an extra measurement at the beginning of the regression line.	161
5.6	The pseudo-3D calibration results in higher attenuation at small distances (compared to two-dimensional calibration).	161

5.7	The two-dimensional calibration results in higher slopes (compared to pseudo-3D calibration).	162
5.8	The pseudo-3D <i>Errors_on_Distance</i> of the anchors are larger than their two-dimensional counterparts.	163
5.9	Boxplot of the position errors with and without (w/o) the preprocessing steps: at the left-hand side for the position errors of the algorithm proposed in this paper and at the right-hand side for the position errors of the minimum mean square error maximum likelihood algorithm. A maximum likelihood algorithm with a mean square error cost function has a lower position error median when our preprocessing is applied.	164
5.10	Comparison of the cdfplot of the two-dimensional and pseudo-3D algorithms	166

List of Tables

1.1	Common applications in which localization and tracking would provide a key advantage	2
1.2	Existing or planned satellite navigation systems	3
1.3	Overview of directional antennas	5
1.4	Wireless technologies for indoor localization versus ranging techniques	12
1.5	Ranging techniques versus design goals	16
2.1	IEEE 802.11b/g channels	40
2.2	IEEE 802.15.4 channels	40
2.3	Power measurement in 2MHz bandwidth (in dBm) of a Wi-Fi transmitter (channel 6) into IEEE 802.15.4 channels	47
2.4	Overview of simulation conditions	48
2.5	Simulated minimum required distance in meter between an IEEE 802.15.4 DSSS receiver and different IEEE interferers (PER=0.10)	48
2.6	Overview of measurement conditions	52
2.7	Minimum required distance in meter between a Wi-Fi interferer and an IEEE 802.15.4 receiver (PER=0.10)	54
3.1	Min-Max algorithm matrix (for the case that all distances are estimated correctly)	90
3.2	Min-Max algorithm matrix with an obstructed path outlier	91
3.3	Min-Max algorithm matrix with a constructive multipath outlier	93
3.4	Outlier test setup for comparison of simple localization algorithms	94
3.5	Simple localization algorithms with three anchors versus design goals.	96
4.1	Schematic description of the preprocessing	114
4.2	Data available for positioning node 4	115
4.3	Data available for positioning node 55.	115
4.4	Data available for positioning node 54	116
4.5	Real distance between target node 4 and the 12 anchors.	116
4.6	Real distance between target node 55 and the 12 anchors.	117
4.7	Real distance between target node 54 and the 12 anchors.	117

4.8	Where are node 4, node 54 and 55 on the regression lines of the anchors?	120
4.9	Theoretical comparison LiReCoFuL and Relative Location Estimation	141
5.1	Definition of the two-dimensional, three-dimensional and pseudo-3D localization	152
5.2	Position error comparison between P3DLiReFLoA and P3D MMSE, both with and without (w/o) the preprocessing.	164
5.3	Comparison of the two-dimensional and pseudo-3D algorithms . .	167

Nomenclature

(\hat{x}_i, \hat{y}_i)	coordinates of the estimated position
(x_i, y_i)	coordinates of the anchors
\hat{d}_i	estimated distance
dBc	decibels relative to the carrier
EOD	<i>Error_on_Distance</i>
<i>Error_on_Distance</i>	two times the estimated standard deviation of the logarithmic distance distribution
<i>intercept</i>	RSSI-measurement at a distance of 1m
n	signal propagation coefficient
PL	Path Loss
RSQ	square of the Pearson correlation coefficient between RSSI and logarithmic distance
<i>slope</i>	sign-sensitive gradient of the RSSI-logarithmic distance line
ACK	ACKnowledgment
AIS	Air Interface Simulator
ANN	Artificial Neural Network
AoA	Angle of Arrival
BLIP	Bluetooth Local Infotainment Point
CCA	Clear Channel Assessment
CDF	Cumulative Distribution Function
CNSS	Compass Navigation Satellite System
CRC	Cyclic Redundancy Check

CRLB	Cramer-Rao Lower Bound
CSS	Chirp Spread Spectrum
CSV	Comma Separated Values
DAQ	Data AcQuisition
DOI	Degree Of Irregularity
DSSS	Direct Sequence Spread Spectrum
DTDoA	Differential Time Differences of Arrival
ED	Energy Detection
EE	Environment Emulator
FCS	Frame Check Sequence
GLONASS	GLObal NAVigation Satellite System
GPIB	General Purpose Interface Bus
GPS	Global Positioning System
GTS	Guaranteed Time Slot
HMM	Hidden Markov Model
I-signal	Inphase signal
IEEE	Institute of Electrical and Electronics Engineers
IP	Internet Protocol
IRNNS	Indian Regional Navigational Satellite System
ISM	Industrial, Scientific and Medical band
JTAG	Joint Test Action Group (developer of IEEE Standard 1149.1-1990)
LiReCoFuL	Linear Regression based Cost Function for Localization
LiReFLoA	Linear Regression based Fast Localization Algorithm
LOESS	LOcal regrESSion
LOS	Line of Sight
LQI	Link Quality Indication
MAC	Media Access Control

MDS	MultiDimensional Scaling
MLH	Maximum LikeliHood
MMSE	Minimum Mean Square Error
NIC	Network Interface Card
NLOS	Non-Line of Sight
O-QPSK	Offset Quadrature Phase Shift Keying
OFDM	Orthogonal Frequency Division Multiplexing
OSI model	Open Systems Interconnection model
P3DLiReFLoA	Pseudo-3Dimensional Linear Regression based Fast Localization Algorithm
PAN	Personal Area Network
PCB	Printed Circuit Board
PDF	Probability Distribution Function
PER	Packet Error Rate
PHR	Physical layer HeadeR
PIFA	Planar Inverted-F Antenna
PLL	Phase Locked Loop
PSD	Power Spectral Density
PSDU	Physical layer Service Data Unit
Q-signal	Quadrature signal
QZS	Quasi-Zenith Satellite System
RBRLE	Reduced Bias Relative Location Estimation
RFID	Radio Frequency IDentification
RIM	Radio Irregularity Model
RLE	Relative Location Estimation
RSSI	Received Signal Strength Indication
RTLS	Real-Time Localization System
SDR	Software Defined Radio

SHR	SyncHRonization
SIR	Signal-to-Interference Ratio
SNR	Signal-to-Noise Ratio
TDoA	Time Difference of Arrival
TML	TransMission Line
ToA	Time of Arrival
ToF	Time of Flight
USB	Universal Serial Bus
UWB	Ultra-WideBand
VCO	Voltage Controlled Oscillator
VNC	Virtual Network Computing
WCL	Weighted Centroid Location
Wi-Fi	Wireless Fidelity
WSN	Wireless Sensor Network

Nederlandstalige samenvatting

–Summary in Dutch–

Voor het enorm potentieel aan locatie-gebaseerde Internet toepassingen in overdekte omgevingen, zijn er andere technologieën dan GPS nodig. Het is dan ook niet verwonderlijk dat zo veel draadloze technologieën, zoals ultrasoon, infrarood, Bluetooth, RFID, Wi-Fi, ToF (Time-of-Flight) met IEEE 802.15.4 DSSS, AoA (Angle-of Arrival) met IEEE 802.15.4 DSSS, IEEE 802.15.4a UWB, IEEE 802.15.4a CCS (Chirp Spread Spectrum) en experimentele 60 GHz concurreren om de ‘ideale oplossing’ te bieden voor contextbewuste toepassingen. Voor een definitieve doorbraak moeten de lokalisatie-algoritmen nauwkeurig, robuust, eenvoudig, energie efficiënt, goedkoop en schaalbaar zijn. Bovendien moeten deze algoritmen een groot gebied bestrijken.

De IEEE 802.15.4 standaard is speciaal ontworpen voor een eenvoudig, low-cost communicatienetwerk, dat draadloze connectiviteit toelaat in toepassingen met beperkt vermogen. De belangrijkste doelstellingen zijn gemak van installatie, betrouwbare gegevensoverdracht, extreem lage kosten en een behoorlijke batterijautonomie, waarbij het protocol toch eenvoudig en flexibel moet blijven. Met 64-bits adressen schaaft de standaard bovendien zeer goed. In vergelijking met Bluetooth, UWB en Wi-Fi, heeft IEEE 802.15.4 DSSS dus niet alleen de laagste protocol complexiteit, maar ook het laagste energieverbruik. Met behulp van complexe timing technieken, presteren IEEE 802.15.4a UWB systemen beter dan IEEE 802.15.4 DSSS systemen op het vlak van nauwkeurigheid. Maar de IEEE 802.15.4 DSSS overtreft de andere technieken in de andere ontwerpdoelstellingen.

Bij het onderzoek van bestaande lokalisatie-algoritmen merkt men een grote discrepantie tussen simulatoren en real-life situaties. Een beter begrip van de fysische laag resulteert bovendien in betere lokalisatie-algoritmen. Er is helaas een tekort aan experimentele resultaten, verkregen uit echte binnenhuistestbedden.

Ons werk focust daarom op experimenteel algoritmeontwerp met een RSSI-gebaseerd (Receiver Strength Signal Indication) draadloos sensor netwerk, op basis van de validatie in het iMinds w-ilab.t real-life test bed. Dit is een uitgebreide draadloze mesh en sensor netwerkinfrastructuur geïnstalleerd in het iMinds kantoorterrein in Gent, met inbegrip van vergaderruimten, klaslokalen, kantoren en gangen. Op de vaste knooppunten hebben we RadioPerf, een specifieke softwaretool, qua functionaliteit vergelijkbaar met de hulpprogramma's iperf, netperf en nuttcp. Voor ons onderzoek zijn deze w-ilab.t proefbank en RadioPerf onmisbaar.

Dit werk presenteert een bottom-up benadering, gericht op het in het ontwerp

houden van de fysische aspecten. Een hoofdstuk gewijd aan de fysische laag is daarom een hebbeding voor dit boek. Het negeren van het specifieke karakter van RSSI-metingen binnen een gebouw resulteert immers niet in goede lokalisatie-algoritmen. De belangrijkste uitdagingen bij deze metingen zijn snelle fading, interferentie, stralingsdiagram van de antenne en multipath fading.

Het gebruik van RadioPerf laat gelukkig een snelle uitmiddeling van de RSSI-metingen toe, zodat het effect van snelle fading effecten genivelleerd wordt. De IEEE 802.15.4 standaard biedt bovendien vele mogelijkheden om interferentie te vermijden. Een zeer breed verontrustende frequentieband wordt echter aangetroffen bij een magnetron oven: deze vult bijna de volledige 2.4 GHz ISM-band. De conventionele voeding van een magnetron oven, zoals de Primo-MG1-B, bestaat gelukkig uit een step-up transformator en een spanningsverdubelaar met een diode gelijkrichter. Dit soort magnetron straalt alleen op de positieve pieken van netvoeding. Onze experimenten tonen aan dat er geen emissie plaatsgrijpt op de negatieve pieken van de netvoeding. Korte pakketten die op deze ogenblikken verzonden worden, leveren dus altijd de juiste RSSI op.

Uitgebreid experimenteel werk presenteert verder de vereiste minimum afstanden tussen een stoorbron en een IEEE 802.15.4 DSSS-ontvanger. Wanneer twee verschillende technologieën ondergebracht zijn in dezelfde behuizing, kan interferentie enkel worden vermeden met behulp van een vermindering van het zendvermogen van de stoorbron en een soort van CCA (clear channel assessment).

De paar dips in het stralingspatroon van de PIFA (planar inverted-F antenne) krijgen onze volledige aandacht. Uit een ander experiment blijkt evenwel dat multipath fading veruit de meest irritante factor is voor binnenhuis lokalisatie.

Deze dips in het antenne stralingspatroon en, nog belangrijker, de constructieve en destructieve multipath fading zijn dus de belangrijkste oorzaken voor de aanwezigheid van uitschieters in de RSSI-metingen. De eerste stap van een lokalisatie-algoritme converteert deze RSSI-metingen naar afstanden. Deze afstandsuitschieters resulteren in een snelle degeneratie van het lokalisatie-algoritme. Het verwijderen van uitschieters is bij lokalisatie helaas zelden aangepakt. Dit eindwerk richt zich dus op de veerkracht tegen uitschieters in veel gebruikte technieken die de individuele berekende afstanden naar een positie converteren. Geometrische multilateratie, statistische multilateratie, maximum likelihood en een min-max lokalisatie-algoritme worden hier onder dezelfde (extreme) multipath fading voorwaarden vergeleken. Uit onze simulaties blijkt dat er een trade-off bestaat tussen complexiteit, nauwkeurigheid en robuustheid tegen uitschieters. Zoals verwacht zijn alle algoritmen kwetsbaar voor de aanwezigheid van uitschieters. Een goede pre-processing techniek is dus nodig om deze uitschieters te elimineren.

In het tweede deel van dit boek, worden eenvoudige lokalisatie-algoritmen ontwikkeld. Extra aandacht wordt besteed om valkuilen te vermijden door het in rekening te brengen van de fysische aspecten, door het opnemen van een goede preprocessing en door met statistische en geometrische hulpmiddelen het effect van uitschieters te verminderen.

Het w-ilab.t real-life test bed is een reeds uitgerold sensor-netwerk met een

groot aantal knooppunten, dat een groot datavolume kan genereren. Het is gangbaar om statistiek te gebruiken voor de analyse van deze (random) gegevens.

Deze thesis maakt geen uitzondering op deze efficiënte manier van werken, die goede resultaten oplevert. Daarom maakt dit werk overvloedig gebruik van waarschijnlijkheidsverdeling, gemiddelde waarden en medianen, lineaire regressie, parametrische en niet-parametrische statistische tests.

Er wordt eerst een statistisch lineair regressiemodel geconstrueerd. Dit eenvoudig model houdt de onderliggende fysische aspecten transparant om foute implementaties te voorkomen. In een eerste stap wordt het model gebruikt om de beste beschikbare ankers dynamisch te selecteren en kalibreren. Een eerste selectiecriteria is de lineariteit van propagatiemodel: de knooppunten met hoogste correlatie tussen de RSSI en de logaritmische afstand worden gepromoveerd tot ankers. Deze ankers worden in de uiteinden van het gebouw aangetroffen, niet in de wandelgangen.

Vervolgens introduceren we de *Error_on_Distance* parameter en we baseren het tweede criterium hierop. Centrale nodes met de kleinste *Error_on_Distance* vervullen de ankerselectie. De *Error_on_Distance* wordt gedefinieerd als twee keer de geschatte standaarddeviatie van de logaritmische afstand frequentieverdeling en wordt verkregen na een verwisselen van de RSSI-as en de logaritmische afstand-as. De regressielijn tussen de RSSI en de logaritmische afstand tussen het anker en alle andere knooppunten bepaalt de (individuele) kalibratie van het anker.

Dan volgt de preprocessing. Eerst worden alle metingen waarbij de berekende afstand groter is dan een limiet (die bepaald wordt door de gevoeligheid van de ontvanger en het zendende energieniveau) als uitschieters beschouwd en verwijderd. De nauwkeurigheid van het model wordt vervolgens gebruikt om slechte RSSI-metingen te elimineren. Na het opschonen van de gegevens kan een snel maximum likelihood algoritme (gebaseerd op *Error_on_Distance*) gebruikt worden om afstandsramingen uit te middelen.

Nu worden verschillende lokalisatie-algoritmen voorgesteld. Het eerste algoritme LiReFLoA gebruikt snelle geometrische principes, die gebaseerd zijn op *Error_on_Distance* en de nauwkeurigheid van het model om een positie te verkrijgen. Nogmaals is er zorg besteed aan het verwijderen van uitschieters die veroorzaakt worden door constructieve multipath fading.

Het tweede nieuwe lokalisatie-algoritme heet LiReCoFuL (Linear Regression based Cost Function for Localization) omdat het een kostenfunctie gebruikt die op lineaire regressie steunt. LiReCoFuL grijpt immers terug naar de basis van een lineair regressiemodel. Net zoals bij de meeste kostenfuncties wordt hierbij homoscedasticiteit verondersteld. Voor LiReCoFuL betekent dit concreet dat de standaarddeviaties van de fouttermen op de logaritmische afstand constant zijn en dus niet afhangen van de RSSI-waarde. Dit is een realistisch uitgangspunt voor de kwaliteitsvolle (zeer goed gecorreleerde) ankers. LiReCoFuL respecteert verder het onderliggende propagatiemodel en maakt gebruik van *Error_on_Distance* om de kostenfunctie op te bouwen.

Deze thesis eindigt met de uitbreiding van 2DLiReFLoA (tweedimensionale

LiReFLoA) tot een pseudo-driedimensionaal lokalisatie-algoritme. Dit algoritme P3DLiReFLoA maakt gebruik van tweedimensionale projectie technieken om een object te vinden in een driedimensionale ruimte. 2DLiReFLoA en P3LiReFLoA delen dezelfde geautomatiseerde tweedimensionale selectie van de ankers.

Voor P3DLiReFLoA is alleen een nieuwe kalibratie nodig om de regressieparameters en *Error_on_Distance* te bekomen. Deze kalibratie kan worden geïnitieerd door de persoon die moet gelokaliseerd worden: onder de geselecteerde ankers triggert hij de gezochte mobiele node. Daarna kan deze persoon overal in het gebouw gevonden worden.

English Summary

As there is a huge potential for location-based Future Internet applications in indoor environments, other technologies than GPS are needed. No wonder that so many wireless technologies, like ultrasonic, infrared, Bluetooth, RFID (Radio-Frequency IDentification), Wi-Fi, ToF (Time-of-Flight) with IEEE 802.15.4 DSSS, Angle-of-Arrival (AoA) with IEEE 802.15.4 DSSS, IEEE 802.15.4a UWB, IEEE 802.15.4a Chirp Spread Spectrum (CSS) and experimental 60 GHz compete to offer the “ideal solution” for context-aware applications. For a final breakthrough, the localization algorithms need to be accurate, robust, simple, energy efficient, cheap and scalable. Furthermore, these algorithms should cover a large area.

The IEEE 802.15.4 standard is designed for a simple, low-cost communication network that allows wireless connectivity in applications with limited power and relaxed throughput requirements. The main objectives are ease of installation, reliable data transfer, short-range operation, extremely low cost, and a reasonable battery life, while maintaining a simple and flexible protocol. Having 64-bit extended addresses, the standard scales extremely well. Furthermore, compared to Bluetooth, UWB and Wi-Fi, IEEE 802.15.4 DSSS has not only the lowest protocol complexity, but also the lowest power consumption. Using complex timing techniques, IEEE 802.15.4a UWB systems perform better than IEEE 802.15.4 DSSS systems on accuracy. The IEEE 802.15.4 DSSS ranging technique, however, definitely outperforms the other techniques in the other design goals.

In present algorithm design, there is a large discrepancy between simulators and real-life situations. Furthermore, a better understanding of the physical layer results in better localization algorithms. Unfortunately, there is a shortage of experimental results, obtained from real indoor test beds.

Therefore, our work focuses on experimental Receiver Strength Signal Indication (RSSI-based) Wireless Sensor Network (WSN) indoor localization based on the validation in the iMinds w-ilab.t real-life test bed. This is an extensive wireless mesh and sensor network infrastructure installed at iMinds office premises in Ghent (Belgium), including meeting rooms, classrooms, offices and corridors. On the fixed nodes we have RadioPerf, a specific software tool, in terms of functionalities very similar to the iperf, netperf and nuttcp tools. For our research this w-ilab.t test bed and RadioPerf have been indispensable.

This work presents a bottom-up approach, aimed at keeping the physical aspects into the design. Therefore, a chapter dedicated to the physical layer is a must-have for this book. Ignoring the specific character of indoor RSSI-measurements does not result in good localization algorithms. The major challenges with in-

door RSSI are fast fading, interference, antenna radiation pattern and multipath fading. Fortunately, the usage of RadioPerf enables a fast averaging of the RSSI-measurements, leveling out the short-time fading effects. The IEEE 802.15.4 standard offers many possibilities to deal with interference. A very wide disturbing frequency band is encountered from a microwave oven, almost filling the complete 2.4 GHz ISM band. The conventional power supply of a microwave oven, like the Primo-MG1-B, consists of a step-up transformer and a diode rectified voltage doubler. This kind of microwave only radiates on the positive peaks of mains supply. Our experiments show that there is no emission on the negative peaks of the mains supply. Therefore, short beacons always deliver the correct RSSI, when transmitted at these instants. Extended experimental work further presents minimum required distances between an interferer and a IEEE 802.15.4 DSS receiver. When two different technologies are housed in the same package, interference can only be avoided using a power reduction of the interferer and a kind of clear channel assessment. The few radiation dips in the inverted-F antenna pattern get our full attention. Another experiment, however reveals that multipath fading is a very annoying factor for indoor localization.

The few antenna radiation pattern dips and, more importantly, the constructive and destructive multipath fading are very important causes for the presence of outliers in the RSSI-measurement. The first step of a localization algorithm converts these RSSI-measurements to distances. These distance outliers results in a quick degeneration of the localization algorithm. Unfortunately, the mitigation of measurement outliers in localization has seldom been addressed. Therefore, this dissertation focuses on the resilience to outliers in widely used techniques that convert the individual calculated distances (ranges) to a position. Geometric multilateration, statistical multilateration, maximum likelihood and a min-max localization algorithm are compared here under the same (extreme) multipath fading conditions. Our simulations show that there is a trade-off between complexity, accuracy, and robustness to outliers. As expected, all algorithms are vulnerable to the presence of outliers. Therefore a good preprocessing technique is needed to eliminate these outliers.

In the second part of this book, simple localization algorithms are developed. Special care is taken to avoid the pitfalls by keeping the physical aspects into the design, by incorporating a good preprocessing and by using statistical and geometric tools in order to decrease the effect of outliers.

The w-ilab.t real-life test bed is a pre-existing sensor network with a large number of nodes, able to generate a large data volume. It is a common practice to use statistical procedures to study this (random) data. This thesis makes no exception on this efficient way of work with good results. Therefore, throughout this work, there is abundant use of probability density functions, mean and median values, linear regression, parametric and non-parametric statistical tests.

First, a statistical model, based on linear regression is constructed. This simple model keeps the underlying physical aspects transparent in order to avoid faulty implementations. In a first step, it is used to dynamically select and calibrate the

best available anchors at a certain time. A first selection criterion is the linearity of the path loss model: the nodes with highest correlation between the RSSI and the logarithmic distance are promoted to anchors. These anchors are found in the extremities of the building, not in the corridors.

Next, we present the *Error_on_Distance* parameter and base the second criterion on it. The *Error_on_Distance* is defined as twice the estimated standard deviation of the logarithmic distance frequency distribution, obtained after an axis swap of the RSSI-axis and the logarithmic distance-axis. The linear regression line between the RSSI and the logarithmic distance between the anchor and all other nodes determines the (individual) calibration of that anchor.

Next, the preprocessing follows. First, all measurements where the calculated distance is greater than a limit (determined by the receiver sensitivity and the sending power level) are considered outliers and dropped. Furthermore, the model accuracy is used to eliminate bad RSSI-measurements. After this data cleaning, a fast maximum likelihood algorithm based on *Error_on_Distance* is used for leveling out the same distance estimates.

Next, different localization algorithms are presented. The first algorithm LiReFLoA (Linear Regression based Fast Localization Algorithm) further uses fast geometric principles, based on the model accuracy and *Error_on_Distance* to obtain a position. Again, care is taken to avoid the presence of the outliers, caused by constructive multipath fading.

The second new localization algorithm is called LiReCoFuL (Linear Regression based Cost Function Localization) because it uses a linear regression based cost function in a maximum likelihood algorithm. LiReCoFuL gets back to the basics of a linear regression model. A basic assumption is homoscedasticity, meaning that the standard deviations of the error terms on the logarithmic distance are constant and do not depend on the RSSI-value. This is a realistic assumption for the high-quality (highly correlated) anchors. LiReCoFuL further respects the underlying physics of the propagation model, and uses *Error_on_Distance* to build the cost function.

This thesis ends with the expansion of 2DLiReFLoA (two-dimensional LiReFLoA) to a pseudo-three-dimensional localization algorithm. P3DLiReFLoA uses two-dimensional projection techniques to find an object in a three-dimensional space. 2DLiReFLoA and P3DLiReFLoA share the same automated two-dimensional selection of the anchors. For the P3DLiReFLoA, only a new calibration is needed to obtain the new regression parameters and *Error_on_Distance*. This calibration can be initiated by the person to be located: beneath the selected anchors, he triggers the mobile target. After this calibration, this person can be located anywhere in the building.

1

Introduction

1.1 Importance of indoor localization

A huge part of the emerging Future Internet applications can significantly benefit from or even is directly dependent on context awareness, in particular location awareness. This information might pertain to the location of the user, to the location of devices / appliances or both of them and hence, enable a lot of powerful applications. Table 1.1 provides an overview of applications (ranging from military to consumer) in which localization and tracking capability would provide a key advantage, partially based on [1]. This table illustrates the huge market potential for localization systems. The emergence of satellite navigation systems has triggered a significant progress in terms of location-based services. Table 1.2 shows an overview of these existing or planned systems. The usage of these services, however, has important limitations with respect to applicability (GPS only works well outdoor), cost, and energy consumption. As there is a huge potential for location-based Future Internet applications in indoor environments, other technologies than GPS are needed. Our work focuses on experimental Receiver Strength Signal Indication (RSSI-based) Wireless Sensor Network (WSN) indoor localization.

User type	Application
Military	Urban combat
	Building clearing
	Ship boarding
	Safe navigation
Law enforcement	Hostage rescue
	Ship boarding
	Crowd control
	Traffic control
Firefighting	Residential and apartment building fires
	Complex building fires
	Ship fires
	Forest fires
	Subterranean rescue operations
	Volcanic eruption detection
Medical	Tracking doctors
	Tracking equipment
	Localization of elderly
Industrial	Tracking and tracing persons
	Production control
	Entrance admittance
	Warehouse management
Civilian users	Detention facilities
	Private security guards
	Transportation of hazardous materials
	Museum guidance
Consumer	Geocaching
	City games (e.g. The Target)
	Social networks (e.g. Foursquare)
	Hiking safety
	Automated Enforcement System detectors

Table 1.1: Common applications in which localization and tracking would provide a key advantage

1.2 State of the art

A good starting point of the study of localization algorithms can be found in [2–6]. There are three interacting phases towards realization of accurate location-based applications: hardware design, ranging and location estimation. Ranging involves the process of the estimation of the distances to a target. Once the distances are known, a position estimation is calculated.

Acronym	Name	Operated for
GPS (NAVSTAR)	Global Position System	USA
GLONASS	GLObal NAVigation Satellite System	Russia
Galileo	(after Italian astronomer)	European Union
Beidou	(after Big Dipper constellation)	China
CNSS	Compass Navigation Satellite System	China
IRNSS	Indian Regional Navigational Satellite System	India
QZSS	Quasi-Zenith Satellite System	Japan

Table 1.2: Existing or planned satellite navigation systems

This section presents the state-of-the-art (indoor) localization systems. First, an overview is given of state-of-the-art ranging techniques and location estimation methods. Next, we discuss how ranging techniques and location estimation methods are applied in combination with different wireless technologies. In the remainder of this chapter, it is shown how this book will advance the highest level of development in localization.

1.2.1 State-of-the-art ranging techniques

Localization methods can be divided into two categories [7]: range-based and range-free. The former is defined by protocols that use absolute point-to-point distance estimates (range) or angle estimates for calculating location. The latter makes no assumption about the availability or validity of such distance or angle information. Range-free methods are found in rather theoretical, not empirical work. [8] compares two range-free localization algorithms. In environments with obstacles, many range-free techniques that have been proposed to improve the localization accuracy are useless [9].

There are different wireless technologies and measurement concepts for range-based indoor localization (see Table 1.4). Commonly proposed techniques to perform ranging are:

- RSSI (Received Signal Strength Indication) is an indication of the power level received by a receiver expressed in dBm. This value is then used to estimate the distance between transmitter and receiver. The physics behind this technology are the power level decay with distance. RSSI is available in most RF receivers.
- ToA (Time of Arrival), also called ToF (Time of Flight) uses the travel time of a radio frequency wave from one transmitter to one receiver. With the

speed of light, the distance is calculated. ToA requires precise synchronization of timers at both transmitter and receiver.

- TDoA (Time Difference of Arrival) is also based on the speed of light. Usually, a multilateration is performed. Here, the position is calculated with at least three spatially separated receiver sites (and one transmitter, being the object to be localized). The difference of the time of arrival at two receivers will narrow the possible position to one half of a two-sheeted hyperboloid. The knowledge of the time of arrival at the third receiver is needed to calculate the unknown position. TDoA only requires precise synchronization at the receivers. In many wireless sensor networks, TDoA is based on the time difference between simultaneously transmitted radio and ultrasound pulses as in the Cricket system [10], as typical WSN clocks are too slow for applying the first approach only.
- AoA (Angle of Arrival) determines the angle of an incident RF wave, which requires directional antennas. An overview of these antennas is given in table 1.3. As can be seen in the first column, there are two main antenna categories: antennas with one driven element and antennas with multiple driven elements.

In the first category there are three principles. A first subcategory is based on the antenna gain. Yagi antennas, parabolic antennas and short backfire antennas are examples of this antenna type. A second subcategory is based on the antenna dip in the radiation pattern. Examples include the loop antenna and the quad antenna. In the third subcategory an antenna consists of a monopole, surrounded by passive elements, which can be individually switched between ground, working as reflectors, or isolation, working as beam directors [11]. An example of this kind of antenna is SPIDA (SICS Parasitic Interference Directional Antenna).

The second category, which has multiple driven elements, consists of two subcategories. In the first subcategory, each driven element has its own receiver. A typical example is a phased array (using the phase between the elements). In the second subcategory, the same receiver is shared between the driven elements. Examples include the Pseudo-Doppler shift antenna, the Watson-Watt antenna and the Correlative interferometer antenna.

- DTDa (Differential Time Differences of Arrival) uses the difference of TDoA-measurements. This is done to overcome the time synchronization of both transmitter and receivers. This is accomplished by introducing a fourth anchor which is responsible for initiating the TDoA-measurement by transmitting a special message. In this way the anchors time offsets can be computed [12].

#Driven elements	Working on	Example
Single	Gain	Yagi Parabola Short backfire
Single	Dip	Loop antenna Quad antenna
Single	Switched parasitics	SPIDA [11]
Multiple (with multiple receivers)	Phase	Phased array
Multiple (with single receiver)	Amplitude and/or Phase	Pseudo-Doppler shift Watson-Watt Correlative interferometer

Table 1.3: Overview of directional antennas

- Proximity uses a very weak sending power, if a message is received, the receiver knows it is the vicinity of the sender. We cannot infer anything, if the message is not received.
- Hybrid techniques: the combination of any of the previous techniques

An unobstructed path between sending and receiving antennas is called a LOS-path. Conversely, when the direct path is blocked, there is a NLOS condition. The latter may result in diffraction, refraction, absorption, scattering or multipath reflection, creating different paths to the receiver. The multiple signals arrive at a receiver at different times and different amplitudes. Therefore, all existing ranging techniques are affected. ToA-based systems, however, seem to be extremely vulnerable to NLOS conditions [13, 14].

1.2.2 State-of-the-art location estimators

Once the ranging measurements between the fixed anchor points (whose positions are already known) and the (mobile) object to be located (whose position is unknown) are available, it is possible to utilize several methods for the estimation of the location of the object:

- Fingerprinting versus non-fingerprinting methods:
 - Typical for a fingerprinting is the use of a large database. During the time consuming (in the order of several days) off-line phase (also called training phase), this database is filled with measurements (e.g. the RSSI-values recorded by nodes knowing their own position). The

online phase is the positioning of a target: here, a new measurement is compared to the values in the database. The stored measurement that is closest to the measurement of the target gives the estimated position. A drawback of this method is that the database needs to be filled with new measurements if the environment changes (e.g. adding a new bed in a hospital localization system) [15, 16].

- Non-fingerprinting methods do not require a database, and these methods are faster. When the environment doesn't change, fingerprinting methods usually give a better accuracy.
- Geometric versus statistical methods.
 - Geometric techniques use geometry to calculate the position from at least three ranging (RSSI/distance, time, angle) measurements. An example is geometric multilateration [17].
 - The use of statistics is widely accepted in location estimation. Here, the frequency distribution of the distances is considered for making an estimation of the position. Mainly, there are three different methods: statistical multilateration [17], maximum likelihood estimators and Bayesian estimators.
 - * In its simplest form statistical multilateration minimizes the sum of the square of the ranging errors (e.g. distance errors). There also exists a weighted least square approach, like in [18]. Here the measured values are first weighted, before the minimization: e.g. high RSSI-values are given a higher weight. In indoor environments this leads to the unjust preference of the paths with the most constructive multipath fading according to our recently performed study [19].
 - * Maximum likelihood methods make use of a cost function. Dependent on the kind of cost function, it needs to be minimized or maximized to find the most likely position. Several cost functions exist. In [20] the simplest and most widely accepted method (minimum mean square error) is presented. Some more cost functions, not only for RSSI but also for ToA-measurements, are presented in [21]. A linear regression based cost function is introduced in previous work [22]. This cost function has less local extremes than the other cost functions. There are two approaches for finding the extremes of a cost function. The first method uses an analytical method, like in [23]. Another method is the grid method where the cost function is calculated in each point of the grid, like in [24]. The first method is faster, but there is a greater chance to get stuck in a local extreme.

- * Bayesian localization methods are based on Bayes' theorem [25] and therefore incorporate some prior knowledge in the estimator. Extensively used methods are Kalman filters, Particle filters [26, 27] and hidden Markov models (HMM) [28]. Another example of a Bayesian method can be found in [29].

1.2.3 State-of-the-art wireless technologies

Ranging techniques and location estimation methods all have their advantages, but also their shortcomings and can be applied to different wireless technologies. A short overview by wireless technology is listed below:

- Ultrasonic localization [30] measures the time difference between an emitted acoustic wave and its reflection. Because ultrasonic waves can't propagate through walls, the infrastructure needs to be deployed in every room. These systems have low coverage. Furthermore, ultrasonic network routing is hard to implement, resulting in the use of combination of ultrasonic and another technology (like IEEE 802.15.4 in [31]). In [4] three ultrasonic positioning systems are reported:
 - Active bats [32] developed by AT&T, whereby a quite sensitive alignment of the large number of receivers is needed.
 - Cricket [10] developed by MIT Laboratories. The system uses both ultrasonic and RF to calculate the distance from the difference in time of arrival between the ultrasonic and RF waves. This system is very sensitive to multipath fading [33]: if the direct path is obstructed, the ultrasonic reflection is used. This results in a too long distance.
 - DOLPHIN [34]. This system has a RF backbone network.

Another problem with ultrasonic systems is that the speed of ultrasound is highly correlated with temperature [33, 35].

- Infrared localization [36] detects a beacon that is sent in the infrared spectrum. Because these signals cannot travel through walls (like ultrasonic signals), infrared systems need room deployment and have low coverage. Furthermore, this coverage is affected by sunlight [37]. An example of an indoor infrared positioning system is Active Badge, developed by AT&T Cambridge [4, 37]. To overcome a complex network routing capability in the infrared region, a telephone line is used here to centralize the different measurements.
- A first example of RF-based localization technology is based on the Bluetooth technology [36]. An example of a Bluetooth system is BlipTrack

(Bluetooth Local Infotainment Point Tracking), developed by Ericsson [38]. Bluetooth not only has a limited range, but also a restricted number of devices [39]. The systems use the received signal strength indication (RSSI) for localization. Installation and deployment require calibration, special hardware, tags, beacons, and previously selected access points to serve as anchors; operation and maintenance can be an issue, as well as cost and scalability of the system and recalibration [17]. A disadvantage of a Bluetooth-based positioning system is that the system can only provide accuracy from 2 m to 3 m with a delay of about 20 s [5]. Furthermore, Bluetooth is power-consuming due to the frequency-hopping spread spectrum technology employed on the physical layer [40]. The upcoming Bluetooth 4.0 standard will have built-in support for indoor positioning [41], but this is still experimental and not available in products yet. Actually, the Bluetooth 4.0 version aims at becoming a low energy wireless technology [42], featuring

- Ultra-low peak, average and idle mode power consumption
- Ability to run for years on standard coin-cell batteries
- Low cost
- Multi-vendor interoperability
- Enhanced ranging

These features have already been found in IEEE 802.15.4 based ZigBee devices for many years.

- Another RF-based technology is RFID localization [36, 43]. RFID localization generally uses RSSI signals. Examples are: SPOT-ON [44] and LAND-MARC [45]. The latter system is revisited by [46]. A more sophisticated RFID system may use the Differential Time of Arrival (DToA) concept. An example is WHERENET [47]. All RFID systems need numerous infrastructure components installed and maintained in the working area of an RFID positioning system [5].
- Most Wi-Fi localization [48–50] systems use RSSI signals, such as EKA-HAU [51], RADAR [15] COMPASS [52] and Aeroscout [53], while fewer Wi-Fi systems apply the Time (Difference) of Arrival (ToA/TDoA) concept [54]. RSSI-measurements on Wi-Fi are tricky since the absolute accuracy on the RSSI is not specified in the standard, making the RSSI specification manufacturer dependent. In addition, due to complex indoor environments (presence of multipath fading and various interfering sources), the performance of the positioning systems is not very accurate (typical accuracies of several meters are reported). Many RSSI-based localization techniques use propagation models for distance estimation, others rely on

fingerprinting techniques in combination with large databases to store the fingerprint information versus location. Consequently, location estimations are complex, costly, and not scalable. The calculation time increases significantly with the number of users [5].

- Instead of Wi-Fi, the cheaper IEEE 802.15.4 DSSS [55] (the original version of IEEE 802.15.4 has been defined in 2003 [56], and revised in 2006 [57] and in 2011 [58]) technology is often used for RSSI-based indoor localization, e.g. Freescales MC13224V ZigBee Platform in Package (PiP) [59]. Unlike Wi-Fi, the RSSI is specified in the IEEE 802.15.4 standard. This is the most widely used radio on today's sensor node platforms. ZigBee, WirelessHART, 6LoWPAN and Glowbal all use IEEE 802.15.4 [60–63]. Most devices use the CC2420 chip from Chipcon with a dynamic range of 100 dB and the datasheet specifies the RSSI linearity (the relationship between RSSI and received input power) within the limits of ± 3 dB [19]. Besides the problems with multipath fading and interfering sources, some authors conclude that the large amount of characterization will make the use of signal strength approaches with low power radios practically impossible [64]. Compared to Bluetooth, UWB and Wi-Fi, however, IEEE 802.15.4 DSSS has not only the lowest protocol complexity, but also the lowest power consumption [65]. Motetrack [66] uses a fingerprinting method: in a time consuming training phase a database is filled with RSSI-measurements, and in the online phase a measurement is matched with these previously stored measurements. In a dynamically changing environment (e.g. changing the position of furniture), the time consuming training phase needs to be redone in order to get accurate results. In [19] we present a fast linear regression based localization method with a fast calibration process. Other statistical methods for indoor localization like maximum likelihood and Bayesian estimators are described in [67]. In [68] a maximum likelihood localization algorithm is presented. Bayesian localization algorithms are presented in [29, 69, 70].
- Time-of-Flight (ToF) with IEEE 802.15.4 DSSS. Time-of-flight based approaches use the round trip time between two nodes to estimate the distance. Many of these efforts rely on special hardware in terms of FPGAs for advanced signal processing or the simultaneous reception of signals on several channels, e.g. those by Lanzisera et al. [71], Pichler et al. [72] and Karalar and Rabaey [73]. Localization using ToF is more accurate with UWB radios due to the physical properties of UWB signals, in particular the large bandwidth (see further). Mazomenos et al. were the first to demonstrate ToF based ranging for sensor networks using CoTS hardware, i.e., standard IEEE 802.15.4 [74]. One of the challenges with using ToF is that one clock tick

of a 8 MHz clock typically found on sensor nodes corresponds to a distance of almost 20 m. (Indeed, during the 125ns clock-pulse an electromagnetic wave travels 37.5 m. This distance needs to be halved, because TOF uses the round trip time, back and forth to the transmitter.) Therefore, one needs to take a lot of measurements and average the clock-ticks to achieve higher accuracy. Mazomenos et al. achieved very good results outside with nodes placed in line-of-sight. Under NLOS conditions, the multipath effect can be substantial which can lead to longer ToF-measurements than required for the actual distance between sender and receiver. SICS researchers presented an estimation method, which compensates for this effect by using the variance between measurements from different channels [75]. This way, they significantly improve the results of Mazomenos in scenarios where nodes are not in line-of-sight.

- Angle-of-Arrival (AoA) with IEEE 802.15.4 DSSS. By measuring the angle of arrival of incoming packets, it is possible to triangulate a receiver's position. This method requires directional antennas such as the SICS Parasitic Interference Directional Antenna (SPIDA) [11, 76]. SPIDA is a cheap directional antenna that can be attached to standard sensor nodes such as the TMote Sky from where it can be directed via simple APIs [77]. In contrast to antenna arrays, SPIDA determines the direction of arrival by computing the amplitude (power) of the incoming signal. For positioning nodes based on AoA, [11] also presented a method based on recursive estimation that computes a new position estimate using a (linear) Kalman filter. Please note that AoA localization algorithms rely on a good orientation of the antennas.
- IEEE 802.15.4a UWB [78–81] is a new RF technology following the IEEE 802.15.4a standard. The previously used UWB IEEE 802.15.3a standard has been redrawn [82]. The IEEE 802.15.4a standard [83] is an amendment to the IEEE 802.15.4-2006 standard [57] and describes two additional physical layers (UWB and CSS). Since 2011 this addendum is included in the IEEE 802.15.4 standard [58]. Only few UWB devices are available (at reasonable cost). At this very moment, only PulsOn 410 is available in large quantities. This device has been released on 28 June 2012, costs around \$2000 and consumes 4.2 Watt [84]. The UWB systems use Angle of Arrival, (AoA), RSSI or Time (Difference) of Arrival (ToA/TDoA) [85]. Examples include Ubisense RTLS [86]. This system uses both DToA and AoA and is also quite expensive [5]. There is a number of approaches using ToF, including those by Gezici et al. and Alsindi et al. [87]. The main problems with UWB localization are: time synchronization, power consumption, and the efficient use of bandwidth [85]. Another main problem of UWB is that the receivers are expensive; the complexity has moved from transmitters

to receivers. This is especially bad for sensor networks which only transmit intermittently. DToA uses the difference of ToA between two different technologies, e.g. UWB and ultrasonic. In realistic crowded environments, DToA localization systems do not perform as well as with RSSI localization systems due to shadowing by walking people [88]. Current state-of-the-art UWB algorithms using cooperation are able to achieve very high accuracy (around 10 cm) but unfortunately suffer from very high complexity and therefore cannot work real-time [89, 90].

- Another new radio of the IEEE 802.15.4a standard is Chirp Spread Spectrum (CSS) [58, 83]. Nanotron RLTS is an example [91]. Here, localization is performed with ToA. When the direct path is obstructed, as in many indoors environments, this results in large errors [92, 93].
- Interest in the utilization of the 60 GHz portion of electromagnetic spectrum was originated by a need for wide bandwidth and interference free frequency bands. The higher attenuation at high frequencies makes this technology ideal for high-density short-range links [94]. Although localization algorithms need to span large ranges and do not need high throughput, there are some initial reports on experimental localization in the 60 GHz band [12]. Complete systems do not exist yet. The quality of the link between the transmitter and the receiver is strongly influenced by the openings (open or closed door) between adjacent rooms. Also the presence of people is important: a human body can easily shadow the significant path [95].
- Finally, previous technologies may be combined [96, 97]. These combinations exploit the synergy between two separate positioning systems, which degrade in different ways. However, this comes at a higher complexity and higher computational cost. [96] is an example where the UWB technology combines the different measuring techniques RSSI and ToA. Also two different technologies are combined in COCKTAIL [97], which is a combination of RFID and RSSI.

Table 1.4 gives an overview of the wireless technologies for indoor localization versus the ranging technologies.

	RSSI	ToA (ToF)	TDofA	DTDoA	AoA	Proximity
Ultrasonic	-	x	x	-	x	obstacle avoidance
Infrared	-	-	-	-	x	x
Bluetooth (IEEE 802.15.1)	x	-	x	-	x	
RFID	x	x	x	-	x	x
Wi-Fi (IEEE 802.11)	x	x	x	x	x	x
IEEE 802.15.4 DSSS	x	x	-	-	x	x
hline IEEE 802.15.4a UWB	x	x	x	-	x	-
IEEE 802.15.4a CCS	-	x	-	-	-	-
60 GHz	-	-	-	x	-	-

Table 1.4: Wireless technologies for indoor localization versus ranging techniques

1.2.4 Statistical tools

Observed data representing a physical phenomenon can be classified as being either deterministic or random. Deterministic data are those that can be described by an explicit mathematical relationship. Data representing a random physical phenomenon cannot be described by an explicit mathematical relationship, because each observation is a unique sample of this data [98]. Due to fading in a propagation path, RSSI-measurements belong to the random data category: no explicit mathematical expression can be written for the time histories produced by this random phenomenon. It is a common practice to use statistical procedures to study the properties of random data. This thesis makes no exception on this efficient way of work with good results. Therefore, throughout this work, there is abundant use of probability density functions, mean and median values, linear regression, parametric and non-parametric statistical tests.

In statistics, an outlier is a measurement that is well outside of the expected range of values, and which is often discarded from the data set. Mainly due to indoor multipath fading, RSSI-measurements deviate from their ideal values. However, the physical aspects behind this deviation cannot be denied! Therefore, outliers will not be discarded in our work. We use statistical and geometric tools to detect the unexpected RSSI-measurements. The main goal is to design localization algorithms dealing with outliers that are due to the underlying physical aspects.

In the robust regression technique [99], the measurements away from the regression line are given a lower weight. This pushes the robust regression line away from the outliers. With robust regression line calibration, the errors will be higher for the outliers (compared to the least squares regression calibration). Therefore, robust regression is not used in this work.

1.3 Design goals for localization algorithms

This section describes a wish list of a perfect localization algorithm. The ideal localization system needs to be accurate, covering a large area, robust, simple, energy efficient, low cost and scalable.

1. Accuracy (and precision).

Accuracy is defined as the degree of closeness of measurements of a quantity to that quantity's actual value. In scientific work on localization, this parameter gets so much attention, that one could assume it is the only criterion of today's localization solutions.

2. Extended range.

An excellent extended range means that the localization algorithm is able to cover a large (maximum) distance. Because range-based systems rely on

the physical principles, all these systems are more or less susceptible to the distance limits of these underlying physics. This parameter complements the accuracy: the accuracy needs to be met in the area of interest.

3. Robustness.

Robustness is the resilience of the system, especially when under stress or when confronted with invalid input. Localization systems are jeopardized by failing anchors (which are nodes knowing their own position), interference, fading (both slow and fast fading, both constructive and destructive multipath fading, delay spread), Doppler effect, thermal fade and even man-in-the-middle attacks. We want to offer accurate localization solutions, not only in quasi-ideal circumstances, but also in various real-life environments under uncontrolled circumstances.

4. Simplicity.

In view of future deployment in commercial applications, localization solutions should be simple, hereby avoiding complex calibration procedures. Many solutions today rely on manual and time-consuming fingerprinting processes during the set-up phase. Such solutions cannot deal with fading and the high dynamics of wireless environments (mobility of people, objects and furniture, changing density of people), requiring a recalibration each time the wireless setting changes. In our work, we focus on quasi real-time approaches that can handle the dynamics the wireless indoor environment. With those approaches set-up time is minimized and recalibration is avoided.

5. Energy efficiency.

This work will pay sufficient attention to the energy efficiency of localization solutions, as the algorithms should also run on resource-constrained devices.

6. Low cost.

In view of commercial deployment, cost is a very important criterion that cannot be neglected, when designing a localization solution. While energy efficient systems help to reduce the total cost of ownership, also cheap, readily available commercial wireless devices are desirable.

7. Scalability.

Scalability is the ability of a system, network, or process, to handle growing amount of work in a capable manner or its ability to be enlarged to accommodate that growth. Scalable localization systems result in a high performance. Indeed, the accuracy of the underlying statistical procedures increases when more nodes are encountered. Furthermore, in a larger scaled network more targets can be localized simultaneously.

In the next section a suitable ranging technique will be chosen based on the constraints of these design goals.

1.4 Why Wireless Sensor Networks (WSN)

In section 1.2 an overview of ranging and location techniques has been presented. This has been supplemented with an overview of wireless technologies. Section 1.3 has added an overview of the design goals of a localization system. These characteristics are combined in table 1.5. This table gives an overview of how the ranging techniques meet the design goals. Ultrasound systems need a high density of sensors to compensate for the poor radio range of ultrasonic signals [100]. Time-based ranging techniques, like IEEE 802.15.4a UWB and IEEE 802.15.4a CSS have the highest accuracy. These systems, however, require complex infrastructure, expensive and energy consuming embedded hardware [100]. Like IEEE 802.11 devices, the IEEE 802.15.4 DSSS based systems have an excellent extended range. This topic will be further discussed in section 3.2.1.

	Accuracy	Extended Range	Robustness	Simplicity	Energy efficiency	Low cost	Scalability
Ultrasonic	😊	😞	😊	😊	😞	😊	😞
Infrared	😊	😞	😊	😊	😞	😊	😞
IEEE 802.15.1	😊	😞	😊	😊	😊	😞	😞
RFID	😊	😞	😊	😊	😞	😊	😊
IEEE 802.11	😊	😊	😊	😊	😊	😊	😊
IEEE 802.15.4 DSSS	😊	😊	😊	😊	😊	😊	😊
IEEE 802.15.4a UWB	😊	😞	😊	😞	😊	😞	😊
IEEE 802.15.4a CCS	😊	😊	😊	😞	😊	😊	😊
60 GHz	😊	😞	😞	😞	😞	😞	😞

Table 1.5: Ranging techniques versus design goals

Compared to Bluetooth, UWB and Wi-Fi, IEEE 802.15.4 DSSS has not only the lowest protocol complexity, but also the lowest power consumption [65]. Furthermore, the IEEE 802.15.4 standard is designed for "... a simple, low-cost communication network that allows wireless connectivity in applications with limited power and relaxed throughput requirements. The main objectives are ease of installation, reliable data transfer, short-range operation, extremely low cost, and a reasonable battery life, while maintaining a simple and flexible protocol." [57]. Having 64-bit extended addresses, the standard scales extremely well. Using complex timing techniques, IEEE 802.15.4a UWB systems perform better than IEEE 802.15.4 DSSS systems on accuracy and robustness. The IEEE 802.15.4 DSSS ranging technique definitely outperforms the other techniques in the other design goals and therefore, it is chosen for.

1.5 Outline and research contributions

In this introduction, it was stated that there is a huge potential for localization-based services. Indeed, the applications range from military to consumer (e.g. tracking and tracing of persons and objects, indoor guiding of persons in complex buildings, offering location-based information, etc.). Furthermore, it has been shown that wireless sensor networks are a valuable candidate to fulfill this need. The main advantage of the usage of WSN is that they are scalable, low cost, energy efficient and simple. For a final breakthrough, however, the following problems need to be solved

- In present algorithm design, there is a large discrepancy between simulators and real-life situations. Therefore, this work presents a new approach for developing a localization algorithm based on the validation in a real-life test bed. Furthermore, this work goes back to the basics and fully exploits the underlying physics of the propagation channel in a realistic model.
- The shortage of experimental results, obtained from real indoor test beds (as outlined in [101]), is well known. Therefore, a large test bed is used throughout this thesis for generating a large data volume. Simple statistical methods are employed to efficiently analyze this data.
- A main requirement for mass deployment of wireless sensor networks and corresponding services is the easy installation and configuration, which is realized by the introduction of self-organizing and auto-configuration mechanisms. For location-based services, the same requirements are valid: the presented approach is based on the automatic selection of anchor nodes for RSSI-based indoor localization and hence avoids complex and time-consuming manual configuration and calibration procedures.

- Table 1.5 shows that the accuracy and the robustness of WSN need to be ameliorated. In chapter 4 we will take special care to improve these two characteristics by the design of preprocessing steps in general and, more specifically, the elimination of outliers.

This thesis is structured as follows.

- A better understanding of the physical layer results in better localization algorithms. Therefore, chapter 2 is dedicated to the hardware and the physical layer. Furthermore, the technological challenges are discussed thoroughly.
- A next step is the presentation of the principles of some existing simple localization algorithms. This is done in chapter 3. Attention is paid to the pitfalls of these common localization algorithms.
- Chapter 4 follows with the design of two-dimensional localization algorithms. In a first step, the best available anchors at a certain time are selected and calibrated. Next, a preprocessing algorithm is presented. The design of two different localization algorithms completes this section.
- Chapter 5 extends a two-dimensional localization algorithm to a pseudo-three-dimensional localization algorithm. Execution times barely change.
- In chapter 6 overall conclusions are drawn.

1.6 Scope of this thesis

A thesis on localization algorithms cannot reasonably cover the whole field. Therefore, this section draws the borders of the scope of this thesis.

- In this thesis we focus on RSSI based IEEE 802.15.4 DSSS WSN. As discussed in section 1.4, this standard fulfills the extended range, simplicity, energy efficiency, low cost and scalability design goals.
- Table 1.5 shows that the accuracy and the robustness of WSN need to be ameliorated. Therefore, we focus on simple preprocessing techniques, based on linear regression.
- This work avoids complex mathematical calculations, the underlying physical aspects are explained using distance circles. Here, we focus on “visualizing” the electromagnetic waves. Bayesian estimators, Fingerprinting methods, Hidden Markov models and Artificial Neural Networks are mentioned, but not implemented because they are computationally complex.

- We assume a deployed sensor network with a large number of nodes (which is a realistic scenario for future dynamic wireless indoor environments) and concentrate on the lower OSI-layers. The data is collected over the wired network part of our real-life test bed. Therefore, the centralized and distributed aspects of the algorithms are not studied.
- Furthermore, this thesis presents new concepts for localization algorithms using the iMinds w-ilab.t test bed.
 - This test bed is synchronized. Furthermore, RSSI is less vulnerable to synchronization errors than time based algorithms. Therefore, there is no in depth study of synchronization.
 - The test bed collects the data. This data is processed on a PC. Therefore, our work does not focus on resource optimization of the mobile node.

1.7 Publications

The research results obtained during this PhD research have been published in scientific journals and presented at a series of international conferences. The following list provides an overview of the publications during my PhD research.

1.7.1 A1: Publications in International Journals

- [1] **F. Vanheel**, J. Verhaevert, E. Laermans, I. Moerman and P. Demeester, “Pseudo-3D RSSI-based WSN Localization Algorithm using Linear Regression”, *Journal on Wireless Communication and Mobile Computing*, Article first published online: 23 August 2013. DOI: 10.1002/wcm.2416
- [2] **F. Vanheel**, J. Verhaevert, E. Laermans, I. Moerman and P. Demeester, “Automated Linear Regression Tools Improve RSSI WSN Localization in Multipath Indoor Environment”, *Eurasip Journal on Wireless Communications and Networks*, special issue on Localization in Mobile Wireless and Sensor Networks, Vol. 2011:38, 2011.

1.7.2 P1: Publications indexed by the ISI Web of Science “Conference Proceedings Citation Index”

- [1] **F. Vanheel**, J. Verhaevert and I. Moerman, “Study on Distance of Interference Sources on Wireless Sensor Network,” *European Microwave Conference (EuMC)*, Amsterdam, The Netherlands, pp. EuMC10-1.1-4, 27-31 October 2008.

1.7.3 C1: Articles in other conference proceedings

- [1] **F. Vanheel**, J. Verhaevert, E. Laermans, I. Moerman and P. Demeester, “A Linear Regression based Cost Function for WSN Localization,” 19th International Conference on Software Telecommunications and Computer Networks (SoftCom 2011), Split - Hvar - Dubrovnik, Croatia, 5 pp., 15-17 September 2011.
- [2] J. Verhaevert, **F. Vanheel** and P. Van Torre, “On the Design of Software and Hardware for a WSN Transmitter”, 16th Annual Symposium of the IEEE/CVT, Louvain-La-Neuve, Belgium, p. 11.1-6, 19 November 2009.

1.7.4 C3: Abstracts in conference proceedings

- [1] **F. Vanheel**, J. Verhaevert, E. Laermans and I. Moerman, “Comparison of 2D RSSI based WSN Multipath Faded Indoor Localization Algorithms,” Ghent University Faculty of Engineering and Architecture PhD Symposium, Ghent, Belgium, pp. 212, 1 December 2010.
- [2] **F. Vanheel**, J. Verhaevert and I. Moerman, “On the Concept of RSSI based Indoor WSN Localization using Statistical Tools,” URSI Forum 2010, Brussels, 18 May 2010.
- [3] **F. Vanheel**, J. Verhaevert and I. Moerman, “Study on Calculating 2D Location using WSN in Multipath Environment”, Ghent University Faculty of Engineering and Architecture PhD Symposium, Ghent, Belgium, 9 December 2009.
- [4] **F. Vanheel**, J. Verhaevert and I. Moerman, “Spectral Interference Study of Wi-Fi on Wireless Sensor Networks”, Ghent University Faculty of Engineering and Architecture PhD Symposium, Ghent, Belgium, pp. 108, 5 December 2007.

References

- [1] J. Rantakokko, P. Händel, M. Fredholm, and F. Marsten-Eklöf. *User Requirements for Localization and Tracking Technology: A Survey of Mission-specific Needs and Constraints*. In IPIN 2010: Proceedings of the International Conference on Indoor Positioning and Indoor Navigation, pages 1–9, Zurich, Switzerland, September 2010.
- [2] M. Bal, W. Shen, and H. Ghenniwa. *Collaborative Signal and Information Processing in Wireless Sensor Networks: a Review*. In SMC 2009: IEEE International Conference on Systems, Man and Cybernetics, pages 3151–3156, San Antonio, TX, USA, October 11–14 2009.
- [3] H. Liu, H. Darabi, P. Banerjee, and J. Liu. *Accuracy of RSS-Based Centroid Localization Algorithms in an Indoor Environment*. Survey of Wireless Indoor Positioning Techniques and Systems. IEEE Transactions on Systems, Man and Cybernetics, – Part C: Applications and Reviews 2007, 37:1067–1080, 2007.
- [4] H. Koyuncu and S. H. Yang. *A Survey of Indoor Positioning and Object Locating Systems*. IJCSNS International Journal of Computer Science and Network Security, 10:5:121–128, May 2010.
- [5] Y Gu, A. Lo, and I. Niemegeers. *A Survey of Indoor Positioning Systems for Wireless Personal Networks*. IEEE Communications Surveys & Tutorials, 11:1:13–31, 2009.
- [6] S. Pandey and P. Agrawal. *A Survey on Localization Techniques for Wireless Networks*. Journal of the Chinese Institute of Engineers, 29:7:1125–1148, 2007.
- [7] T. He, C. Huang, B. M. Blum, J. A. Stankovic, and T. Abdelzaher. *Range-Free Localization Schemes for Large Scale Sensor Networks*. In MobiCom 2003, San Diego, CA, USA, September 14–19 2003.
- [8] M. Keshtgary, M. Fasihiy, and Z. Ronaghi. *Performance Evaluation of Hop-based Range-free Localization Methods in Wireless Sensor Networks*. International Scholarly Research Network ISRN Communications and Networking, 2011, 2011.
- [9] S. Zhang, Y. Zeng, L. Chen, D. Chen, and Li Xie. *Performance Evaluation of Localization Algorithms for Mobile Sensor Networks*. Journal of Software 2011, 22:7:1597–1611, 2011.
- [10] N.B. Priyantha. *The Cricket Indoor Location System*. PhD thesis, Massachusetts Institute of Technology, June 2005.

- [11] M. Nilsson. *Localization using Directional Antennas and Recursive Estimation*. In WPNC: Fifth Workshop on Positioning, Navigation and Communication, Hannover, Germany, March 2008.
- [12] F. Winkler, E. Fischer, E. Grass, and G. Fischer. *A 60 GHz OFDM Indoor Localization System Based on DTDOA*. In 14th IST Mobile & Wireless Communications Summit, Dresden, Germany, June 2005.
- [13] W. Wang, J. Y. Xiong, and Z. L. Zhu. *A New NLOS Error Mitigation Algorithm in Location Estimation*. IEEE Transactions on Vehicular Technology, 54:6:2048–2053, November 2005.
- [14] A. Maali, A. Ouldali, H. Mimoun, and G. Baudoin. *Evaluation of UWB Localization under Non Line-Of-Sight (NLOS) Propagation*. In ISWPC 2008: Third International Symposium on Wireless Pervasive Computing, pages 379–382, Santorini, Greece, May 7–9 2008.
- [15] P. Bahl and V. Padmanabhan. *RADAR: An In-building RF-based User Location and Tracking System*. In Proceedings of the IEEE INFOCOM, volume 2, pages 775–784, Tel-Aviv, Israel, March 2000.
- [16] M. Azizyan, I. Constandache, and R. R. Choudhury. *SurroundSense: Mobile Phone Localization via Ambience Fingerprinting*. In MobiCom 2009, Beijing, China, September 20–25 2009.
- [17] D. Munoz, F. Bouchereau, C. Vargas, and R. Enriquez. *Position Location Techniques and Applications*. Academic Press, Burlington, MA, USA, 2009.
- [18] X. An, J. Wang, V. R. Prasad, and I. G. M. M. Niemegeers. *OPT: Online Person Tracking for Context-awareness in Wireless Personal Network*. In Proceedings of the Second International Workshop on Multihop Ad hoc Networks: from theory to reality, pages 47–54, Florence, Italy, May 2006.
- [19] F. Vanheel, J. Verhaever, E. Laermans, I. Moerman, and P. Demeester. *Automated Linear Regression Tools Improve RSSI WSN Localization in Multipath Indoor Environment*. EURASIP Journal on Wireless Communications and Networking 2011, Special Issue: Localization in Mobile Wireless and Sensor Networks, 2011:38, July 2011.
- [20] Y. M. Kwon, K. Mechitoch, S. Sundresh, W. Kim, and G. Agha. *Resilient Localization for Sensor Networks in Outdoor Environments*. In ICDCS 2005: Proceedings of the 25th IEEE International Conference on Distributed Computing Systems, pages 643–652, Columbus, OH, USA, June 2005.

- [21] N. Patwari, A. O. Hero, M. Perkins, N. S. Correal, and R. J. O'Dea. *Relative Location Estimation in Wireless Sensor Networks*. IEEE Transactions on Signal Processing, 51:2137–2248, 2003.
- [22] F. Vanheel, J. Verhaevert, E. Laermans, I. Moerman, and P. Demeester. *A Linear Regression Based Cost Function for WSN Localization*. In SOFTCOM: Proceedings of the 19th International Conference on Software, Telecommunications and Computer Networks, Split, Croatia, September 15–17 2011.
- [23] J. L. Nazareth. *Conjugate-gradient Methods*. Wiley Interdisciplinary Reviews: Computational Statistics, 1:348–353, 2009.
- [24] S. M. Kay. *Fundamentals of Statistical Signal Processing, Volume 1: Estimation theory*. Prentice Hall, Upper Saddle River, NJ, USA, 1993.
- [25] N.J. Gordon, D.J. Salmon, and A.F.M. Smith. *Novel Approach to Nonlinear non-Gaussian Bayesian State Estimation*. IEE proceedings–F, 140:2:107–113, April 1993.
- [26] R.E. Kalman. *A New Approach to Linear Filtering and Prediction Problems*. Transactions of the ASME-Journal of Basic Engineering, 82:D:35–45, 1960.
- [27] J. Biagioni, T. Gerlich, T. Merrifield, and J. Eriksson. *EasyTracker: Automatic Transit Tracking, Mapping, and Arrival Time Prediction Using Smartphones*. In SenSys 2011, Seattle, WA, USA, November 1–4 2011.
- [28] C. Morelli, M. Nicoli, V. Rampa, and U. Spagnolini. *Hidden Markov Models for Radio Localization in Mixed LOS/NLOS Conditions*. IEEE Transactions on Signal Processing, 55:4:1525–1542, April 2007.
- [29] D. Madigan, E. Elnahrawy, R. P. Martin, W.-H. Ju, P. Krishnan, and A. S. Krishnakumar. *Bayesian Indoor Positioning Systems*. In Proceedings of IEEE Infocom, Miami, FL, USA, March 2005.
- [30] E. O. Dijk. *Indoor Ultrasonic Position Estimation Using a Single Base Station*. PhD thesis, Technische Universiteit Eindhoven, 2004.
- [31] S.J. Kim and B. K. Kim. *Accurate Hybrid Global Self-Localization Algorithm for Indoor Mobile Robots With Two-Dimensional Isotropic Ultrasonic Receivers*. IEEE Transactions on Instrumentation and Measurement, 60:10:3391–3404, 2011.
- [32] M. Hazas and A. Hopper. *Broadband Ultrasonic Location System for Improved Indoor Positioning*. IEEE Transactions on mobile Computing, 5:5:536–547, May 2006.

- [33] R. Mautz. *The Challenges of Indoor Environments and Specification on some Alternative Positioning Systems*. In WPNC 2009: Proceedings of the 6th Workshop on Positioning, Navigation and Communication, Hannover, Germany, March 2009.
- [34] Y. Fukuju, M. Minami, H. Morikawa, and T. Aoyama. *DOLPHIN: an Autonomous Indoor Positioning System in Ubiquitous Computing Environment*. In IEEE Workshop on Software Technologies for Future Embedded Systems, pages 53–56, Hakodate, Japan, May 2003.
- [35] Howard W Sams & Co Engineers. *Reference Data for Radio Engineers*. Sams, Indianapolis, IN, USA, 1975.
- [36] J. Hallberg and M. Nilsson. *Positioning with Bluetooth, IrDA and RFID*. Master’s thesis, Lulea University of Technology, Sweden, March 2002.
- [37] R. Want, A. Hopper, V. Falcao, and J. Gibbons. *The Active Badge Location System*. ACM Transactions on Information systems, 40:1:91–102, January 1992.
- [38] *Bluetooth Technology & Products: BlipTrack Traffic*. <http://www.bliptrack.com/traffic/area-of-operations/bluetooth-technology-privacy/>.
- [39] *IEEE Standard for Information Technology – Telecommunications and Information Exchange between Systems – Local and Metropolitan Area Networks – Specific Requirements – Part 15.1: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Wireless Personal Area Networks (WPANs)*. <http://standards.ieee.org/getieee802/download/802.15.1-2005.pdf>, 2005. IEEE 802.15.1 Std.
- [40] M. S. Obaidat, A. Anpalagan, and I. Woungan. *Handbook of Green Information and Communication Systems*. Academic Press, Oxford, UK, 2012.
- [41] M. Guim. *Nokia Research Shows off Indoor Mapping*. <http://thenokiablog.com/2011/11/29/nokia-research-indoor-mapping/>.
- [42] *Bluetooth SIG opens Qualification Program for Bluetooth Core Specification Version 4.0*. <http://www.bluetooth.com/Pages/Press-Releases-Detail.aspx?ItemID=106>, 2010. Press Release.

- [43] B.F. Rolfe, S.W. Ekanayake, P.N. Pathirana, and M. Palaniswami. *Localization with Orientation using RSSI Measurements: RF Map Based Approach*. In ISSNIP: Proceedings of the Third International Conference on Intelligent Sensors, Sensor Networks and Information, pages 311–316, Melbourne, Australia, December 2007.
- [44] J. Hightower, R. Want, and G. Borriello. *SpotON: An indoor 3D Location Sensing Technology Based on RF Signal Strength*. Technical Report UW CSE00-02-02, University of Washington, February 2000.
- [45] L. M. Ni, Y. Liu, I. C. Lau, and A. P. Patil. *LANDMARC: Indoor Location Sensing Using Active RFID*. In Proceedings of the First IEEE International Conference on Pervasive Computing and Communications, pages 407–416, Forth Worth, USA, 2003.
- [46] X. Yinggang, K. JiaoLi, W. ZhiLiang, and Z. Shanshan. *Indoor Location Technology and its Applications Base on Improved LANDMARC Algorithm*. In CCDC 2011: Control and Decision Conference, Mianyang, China, May 23–25 2011.
- [47] *WhereNet Technology ISO/IEC 24730-2*. <http://www.zebra.com/us/en/solutions/technology-need/wherenet.html>.
- [48] S. Mazuelas, A. Bahillo, R. M. Lorenzo, P. Fernandez, F. A. Lago, E. Garcia, J. Blas, and E. J. Abril. *Robust Indoor Positioning Provided by Real-time RSSI Values in Unmodified WLAN Networks*. IEEE Journal of Selected Topics in Signal Processing, 3: 5:821–831, October 2009.
- [49] C. Tseng and S. Cheng. *Location Management Scheme with WLAN Positioning Algorithm for Integrated Wireless Networks*. COMCOM: Computer Communications, 31:18:4304–4311, December 2008.
- [50] A. Kushki, K. N. Plataniotis, and A. N. Venetsanopoulos. *Indoor Positioning with Wireless Local Area Networks (WLAN)*, 2008.
- [51] *EkaHau*. <http://www.ekahau.com/>.
- [52] T. King, S. Kopf, T. Haenselmann, C. Lubberger, and W. Effelsberg. *COMPASS: A Probabilistic Indoor Positioning System Based on 802.11 and Digital Compasses*. In WinTECH: Proceedings First ACM Intl Workshop on Wireless Network Testbeds, Experimental evaluation and CHaracterization, Los Angeles, CA, USA, September 2006.
- [53] *Infrastructure*. <http://www.aeroscout.com/infrastructure>.

- [54] X. Li, K. Pahlavan, M. Latva-aho, and M. Ylianttila. *Comparison of Indoor Geolocation Methods in DSSS and OFDM Wireless LAN Systems*. In VTC 2000 Fall: Proceedings of the 52th Vehicular Technology Conference, pages 3015–3020, Boston, MA, USA, 2000.
- [55] M. Sugano, T. Kawazoe, Y. Ohta, and M. Murate. *Indoor Localization System Using RSSI Measurements of Wireless Sensor Network Based on the ZigBee Standard*. In Proceedings of the Sixth Lasted International Multi-Conference on Wireless and Optical Communications, pages 503–508, Banff, Canada, July 2006.
- [56] *IEEE Standard for Information Technology – Telecommunications and Information Exchange between Systems – Local and Metropolitan Area Networks – Specific Requirements – Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LR-WPANs)*. <http://standards.ieee.org/getieee802/download/802.15.4-2003.pdf>, 2003. IEEE802.15.4-2003 Std.
- [57] *IEEE Standard for Information Technology – Telecommunications and Information Exchange between Systems – Local and Metropolitan Area Networks – Specific Requirements – Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs)*. <http://standards.ieee.org/getieee802/download/802.15.4-2006.pdf>, 2006. IEEE802.15.4-2006 Std.
- [58] *IEEE Standard for Local and Metropolitan Area Networks – Part 15.4: Low-Rate Wireless Personal Area Networks (LR-WPANs)*. <http://standards.ieee.org/getieee802/download/802.15.4-2011.pdf>, 2011. IEEE802.15.4-2011 Std.
- [59] O. Hernandez, V. Jain, S. Chakravarty, and P. Bhargava. *Position Location Monitoring Using IEEE802.15.4 ZigBee Technology*. <http://www.freescale.com/files/microcontrollers/doc/brochure/PositionLocationMonitoring.pdf>.
- [60] *ZigBee Specification Overview*. <http://zigbee.org/Specifications/ZigBee/Overview.aspx>.
- [61] *WirelessHart Overview*. http://www.hartcomm.org/protocol/wihart/wireless_overview.html.
- [62] *IPv6 over Low Power WPAN*. <http://datatracker.ietf.org/wg/6lowpan/charter/>.

- [63] *Glowbal IP: An Adaptive and Transparent IPv6 Integration in the Internet of Things*. <http://iospress.metapress.com/content/x611r3t20n171102>.
- [64] D. Lymberopoulos, Q. Lindsey, and A. Savvides. *An Empirical Characterization of Radio Signal Strength Variability in 3-D IEEE 802.15.4 Networks Using Monopole Antennas*. In *Proceedings of the Third European Workshop on Wireless Sensor Networks*, pages 326–341, Zurich, Switzerland, February 2006.
- [65] J. S. Lee, Y. W. Su, and C. C. Shen. *A Comparative Study of Wireless Protocols: Bluetooth, UWB, ZigBee, and Wi-Fi*. *IECON 2007: Proceedings of the 33rd Annual Conference of the IEEE Industrial Electronics Society*, 1-3:46–51, 2007.
- [66] K. Lorincz and M. Welsh. *MoteTrack: A Robust, Decentralized Approach to RF-Based Location Tracking*. In *LoCA 2005: Proceedings of the International Workshop on Location and Context-Awareness*, Oberpfaffenhofen, Germany, May 2005.
- [67] F. Seco, A. R. Jimenez, C. Prieto, J. Roa, and K. Koutsou. *A Survey of Mathematical Methods for Indoor Localization*. In *WISP 2009: 6th IEEE International Symposium on Intelligent Signal Processing*, pages 9–14, Budapest, Hungary, August 26–28 2009.
- [68] T. Roos, P. Myllymaki, and H. Tirri. *A Statistical Modeling Approach to Location Estimation*. *IEEE Transactions on Mobile Computing*, 1: 1:59–69, January–March 2002.
- [69] G. Chandrasekaran, M. A. Ergin, J. Yang, Y. Chen S. Liu, M. Gruteser, and R. P. Martin. *Empirical Evaluation of the Limits on Localization Using Signal Strength*. In *SECON 2009: Proceedings of the 6th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, pages 333–341, Rome, Italy, June 2009.
- [70] M. Nicoli, C. Morelli, V. Rampa, and U. Spagnolini. *HMM-based Tracking of Moving Terminals in Dense Multipath Indoor Environments*. In *EU-SIPCO 2005: Proceedings of the EURASIP European Signal Processing Conference*, Antalya, Turkey, September 2005.
- [71] S. Lanzisera, D.T. Lin, and K.S.J. Pister. *RF Time of Flight Ranging for Wireless Sensor Network Localization*. In *International Workshop on Intelligent Solutions in Embedded Systems*, pages 1–12, Vienna, Austria, June 2006.

- [72] M. Pichler, S. Schwarzer, A. Stelzer, and M. Vossiek. *Multi-channel Distance Measurement with IEEE802.15.4 (ZigBee) Devices*. IEEE JSAC, 3:5:845–859, 2009.
- [73] T.C. Karalar and J. Rabaey. *An RF TOF Based Ranging Implementation for Sensor Networks*. In IEEE ICC 2006: IEEE International Conference on Communication, Istanbul, Turkey, June 2006.
- [74] E. Mazomenos, D. De Jager, J. Reeve, and N. White. *A Two-Way Time of Flight Ranging Scheme for Wireless Sensor Networks*. In European Conference on Wireless Sensor Networks, Bonn, Germany, February 2011.
- [75] P. Pettinato, N. Wiström, J. Eriksson, and T. Voigt. *Multi-Channel Two-way Time of Flight Sensor Network Ranging*. In European Conference on Wireless Sensor Networks, Trento, Italy, February 2012.
- [76] M. Nilsson. *Directional Antennas for Wireless Sensor Networks*. In Ninth Scandinavian Workshop on Wireless Adhoc Networks, pages 1–4, Uppsala, Sweden, May 4–5 2009.
- [77] E. Öström, L. Mottola, and T. Voigt. *Evaluation of an Electronically Switched Directional Antenna for Real-world Low-power Wireless Networks*. In REALWSN: Third International Workshop on Real-World Wireless Sensor Networks, Colombo, Sri Lanka, December 2010.
- [78] Z. Sahinoglu, S. Gezici, and I. Güvenc. *Ultra-wideband Positioning Systems Theoretical Limits, Ranging Algorithms, and Protocols*. DoCoMo Communications Laboratories USA, CA, USA, September 2008.
- [79] K. Yu and I. Oppermann. *UWB Positioning for Wireless Embedded Networks*. In RAWCON: Proceedings of the IEEE Radio and Wireless Conference, pages 459–462, Atlanta, GA, USA, September 19–22 2004.
- [80] N. Alsindi and K. Pahlavan. *Cooperative Localization Bounds for Indoor Ultra-Wideband Wireless Sensor Networks*. EURASIP Journal on Advances in Signal Processing, 2008:1–13, 2008.
- [81] S. Gezici and et al. *Localization via Ultra-wideband Radios: a Look at Positioning Aspects for Future Sensor Networks*. Signal Processing Magazine, IEEE, 22:4:70–84, 2005.
- [82] *Withdrawal of the 802.15.3a*. <http://standards.ieee.org/about/sasb/nescom/projects/802-15-3a.pdf>.
- [83] *IEEE Standard for Information Technology – Telecommunications and Information Exchange between Systems – Local and Metropolitan Area*

Networks – Specific Requirements – Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs), Amendment 1: Add Alternate PHYs. <http://standards.ieee.org/getieee802/download/802.15.4a-2007.pdf>, 2007. IEEE802.15.4.a-2007 Std. (addendum).

- [84] *PulseOn P410 Datasheet.* <http://www.timedomain.com/p400.php>.
- [85] Z. Xiao, Y. Hei, Q. Yu, and K. Yi. *A Survey on Impulse-radio UWB Localization*, 2010. presented at SCIENCE CHINA Information Sciences.
- [86] *Real-time locating system.* <http://www.ubisense.net/en/rtls-solutions/research-packages.html>.
- [87] N. Alsindi, B. Alavi, and K. Pahlavan. *Measurement and Modeling of Ultrawide-band TOA-based Ranging in Indoor Multipath Environments.* IEEE Transactions on Vehicular Technology, 58:3:1046–1058, March 2009.
- [88] S. Hara and D. Anzai. *Experimental Performance Comparison of RSSI- and TDOA-based Location Estimation Methods.* In Proceedings of the IEEE Vehicular Technology Conference, pages 2651–2655, Singapore, May 2008.
- [89] *Homepage of Samuel Van de Velde.* <http://telin.ugent.be/~slvdveld/>.
- [90] S. Van de Velde, H. Wymeersch, P. Meissner, W. Klaus, and H. Steendam. *Cooperative Multipath-Aided Indoor Localization.* In WCNC 2012: Proceedings of the IEEE Wireless Communications and Networking Conference, pages 3107–3111, Paris, France, April 2012.
- [91] *Nanotron Technologies.* http://www.nanotron.com/EN/PR_nPAL.php.
- [92] J. Wang, Q. Gao, Y. Yu, H. Wang, and M. Jin. *Toward Robust Indoor Localization Based on Bayesian Filter Using Chirp-Spread-Spectrum Ranging.* IEEE Transactions on Industrial Electronics, 59:3:1622–1629, March 2012.
- [93] A. Van Nieuwenhuyse, J. Wyffels, J. P. Goemaere, L. De Strycker, and B. Nauwelaers. *Time of Arrival Based on Chirp Pulses as a means to Perform Localization in IEEE 802.15.4a Wireless Sensor Networks.* Advances in Electrical and Computer Engineering, 10:2:65–70, 2010.
- [94] H. Yang, P. F. M. Smulders, and M. H. A. J. Herben. *Channel Characteristics and Transmission Performance for Various Channel Configurations*

- at 60 GHz*. Eurasip Journal on Wireless Communications and Networking, 2007:19613, March 2007.
- [95] S. Collonge, G. Zaharia, and G. El Zein. *Wideband and Dynamic Characterization of the 60 GHz Indoor Radio Propagation - Future Home WLAN Architecture*. Annales de Télécommunications, Special Issue on WLAN, 58:3-4:417–447, March–April 2003.
- [96] C. D. Wann and H. C. Chin. *Hybrid ToA/RSSI Wireless Location with Unconstrained Nonlinear Optimization for Indoor UWB Channels*. IEEE Wireless Communication & Networking Conference, 1-9:3943–3948, 2007.
- [97] D. Zhang, Y. Yang, D. Cheng, S. Liu, and L. M. Ni. *COCKTAIL: an RF-based Hybrid Approach for Indoor Localization*. In ICC 2010: IEEE International Conference on Communications, Cape Town, South Africa, May 23–27 2010.
- [98] J. S. Bendat and A. G. Piersol. *Random Data, Analysis and Measurement Procedures*. John Wiley & Sons Academic Press, New York, USA, 1986.
- [99] S. Heritier, E. Cantoni, S. Copt, and M. Victoria-Feser. *Robust Methods in Biostatistics*. Wiley, Chichester, UK, 2009.
- [100] P. A. Vicaire and J. A. Stankovic. *Elastic Localization*. Technical Report CS-2004-35, University of Virginia, 2004.
- [101] P. Pivato, L. Palopoli, and D. Petri. *Accuracy of RSS-Based Centroid Localization Algorithms in an Indoor Environment*. IEEE Transactions on Instrumentation and Measurement, 60:3451–3460, 2011.

2

Physical Layer

2.1 Introduction

A better understanding of the physical layer results in better localization algorithms. Therefore, a chapter dedicated to the hardware is a must-have for this book. Section 1.4 revealed that the combination of RSSI as a ranging technique and WSN as a wireless technology possesses the highest potential to meet the design goals for localization that were outlined in section 1.3. This chapter starts with the presentation of a RSSI-based WSN test bed and continues with a study on the RSSI-based technological challenges. It further describes the design of a software designed transmitter, in order to access the physical layer and understand the IEEE 802.15.4 DSSS standard better. This knowledge is needed to develop good localization algorithms.

2.2 The test bed: WSN and RadioPerf

The iMinds iLab.t Wireless Lab or w-ilab.t is an extensive wireless mesh and sensor network infrastructure installed at iMinds office premises in Ghent (Belgium), including meeting rooms, classrooms, offices and corridors. The w-ilab.t is integrated in the CREW platform [1], which is an open federated test platform that facilitates experimentally-driven research on advanced spectrum sensing, cognitive radio and cognitive networking strategies in view of horizontal and vertical spectrum sharing in licensed and unlicensed bands. The w-ilab.t has more than

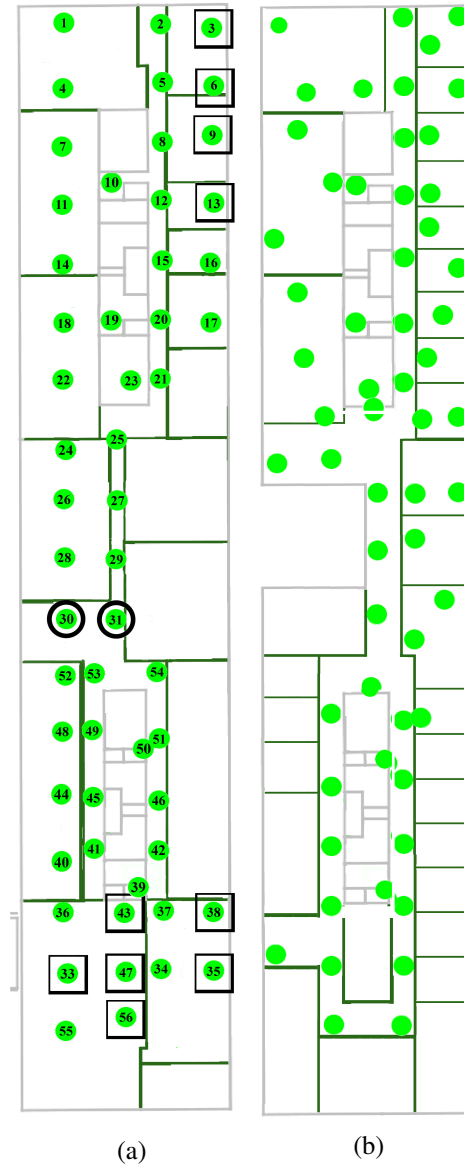


Figure 2.1: Position of sensor nodes on the third floor (a) and second floor (b) of the iMinds office building. Drywall walls are presented by dark green solid lines. The solid gray lines are concrete wall. The black squares and black circles are the selected anchors (see section 4).

200 TMote Sky nodes equipped with a Chipcon CC2420 radio chip operating in the 2.4 GHz frequency band. The nodes are spread over three floors on a grid measuring 12 m by 80 m. The dimensions of the third floor are 16.8 m by 90.0 m or 1512 m². On the third floor there are 56 nodes. All nodes are mounted at the same height (0.15 m beneath a highly conducting ceiling). The central internal walls are constructed of drywall, while walls in the vicinity of the staircases and elevator are made of concrete. There is a microwave oven in each of these concrete zones. At the elevator a lot of metal is used. One side of the longest corridors is concrete and the other drywall. In figure 2.1 (a) this third floor is shown with the position of the nodes and its four narrow but long corridors. On the second floor, there are 58 nodes. The floor plan and the position of the nodes can be found in figure 2.1 (b). The floor is rectangular shaped, but in the center of the floor, there are also outside walls, almost cutting the floor in two smaller rectangles. There is a red brick external wall in front of the concrete outside bearing walls. In this construction, there are many coated aluminium framed windows.

Every node is connected to a corresponding environment emulator (EE). This component can measure real-time energy consumption, emulate battery depletion and generate artificial sensor data, hence emulating real-world application scenarios. The EE is connected to an intermediate node or iNode. The iNodes are further connected to a central management server via Ethernet and are responsible for the configuration and control of the sensor nodes. The iNodes are also connected with 2 IEEE 802.11 (a/b/g) radios and allow Wi-Fi testing. In this work we will only focus on the sensor nodes. Power is fed over Ethernet through Power over Ethernet (PoE) network switches. The control software supports both Tiny OS 1 and Tiny OS 2. The iNodes can be synchronized with a time server. A time accuracy of 2 μ s is obtained and used for data logging. More about our test bed can be found in [2, 3], more about the environment emulator can be found in [4, 5].

On the iNode we have a software tool called RadioPerf which is in terms of functionalities very similar to the iperf, netperf and nuttcp tools [6, 7]. This tool includes a packet generator and Java based configuration and visualization of measurements reported to the central data server. For this study the following items are collected in a log file: the number of packets sent, the number of packets sent with an error, the minimum estimated noise floor, the average estimated noise floor, the maximum estimated noise floor, an identifier of the node generating the packets, the number of packets received, the number of packets lost, the minimum link quality indication (LQI), the average LQI, the maximum LQI, the minimum RSSI, the average RSSI, the maximum RSSI, a time stamp, the node reporting these values and a report sequence number. A screenshot of the RadioPerf tool is presented in figure 2.2. At the upper left-hand side the configuration is set. The highlighted ConfigRadio tab shows the selection of channel 26 (2480 MHz) and the power level of 0 dBm. The lower left-hand side offers a selection of the items to be dis-

played. These items appear at the right-hand side. The highlighted tab represents the number of packets received. This example starts with a good reception of the packets. At the time stamp A the microwave oven is switched on, resulting in a disturbed reception. At the time stamp B the oven is switched off and a good reception follows again.

For our research this w-ilab.t test bed has been indispensable.

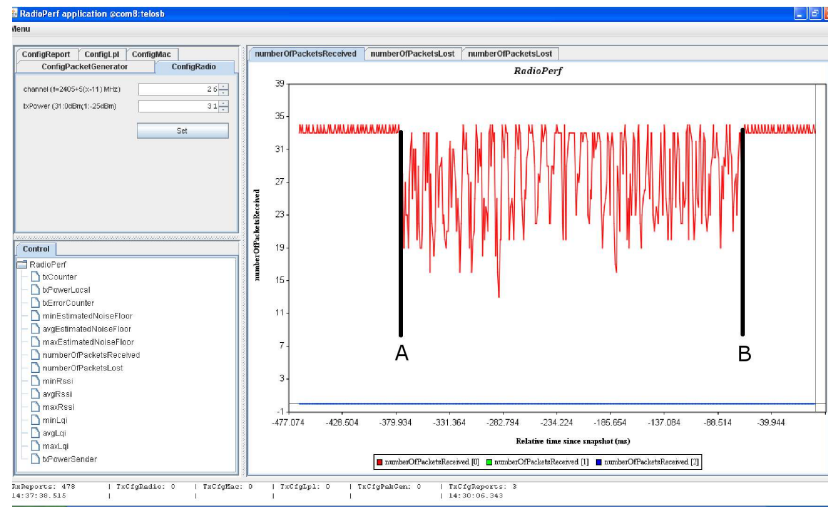


Figure 2.2: RadioPerf with its configuration tabs, selection region and display area.

2.3 Challenges with indoor RSSI

Table 1.4 revealed that both the accuracy and the robustness of IEEE 802.15.4 DSSS (RSSI-based WSN) localization systems can be improved. For achieving these goals, this chapter goes back to the underlying physical aspects to consider. In a first subsection the effect fading is categorized. Next, multipath fading is treated empirically with the help of the iMinds w-ilab.t test bed. It is shown that multipath fading is a very annoying factor for indoor localization. This section finishes with an extended study on spectral interference.

2.3.1 Slow fading versus Fast fading

Using radio waves as the network medium poses several challenges. A first challenge is fading: this is a gradual appearance (or disappearance) of the radio signal. A complete discussion of this topic is beyond the scope of this book. This can be found in [8, 9]. Depending on the coherence time of the channel, one can observe either slow fading or fast fading. The fast fading can be averaged out.

In order to study the effect of time fading in an indoor environment, a test was set up using the w-ilab.t test bed, including all nodes on the third floor. Every node broadcasts 240 packets. Every packet consists of 100 bytes, the inter packet delay is 25 ms. This brief recovery time between packets allows devices to prepare for reception of the next packet. The transmission of one broadcast takes 25 ms times 240 or 6 seconds. Every second, available data was sent to our log file. This implies that some averaging of RSSI-measurements is already done on the mobile nodes. This is done for limiting the amount of reports, while still obtaining a sufficient long measuring time: the receiver of the 240 packets typically needs to send 6 (averaged) RSSI reports. (It could be less if many packets are lost. It could also be one more if sender and receiver are not completely synchronized.) These RSSI-measurements are now read and processed in Matlab on the positioning PC. First, an overall average RSSI is calculated for each sending-receiving pair. Next, this average is abstracted from the minimum value in the corresponding packet group.

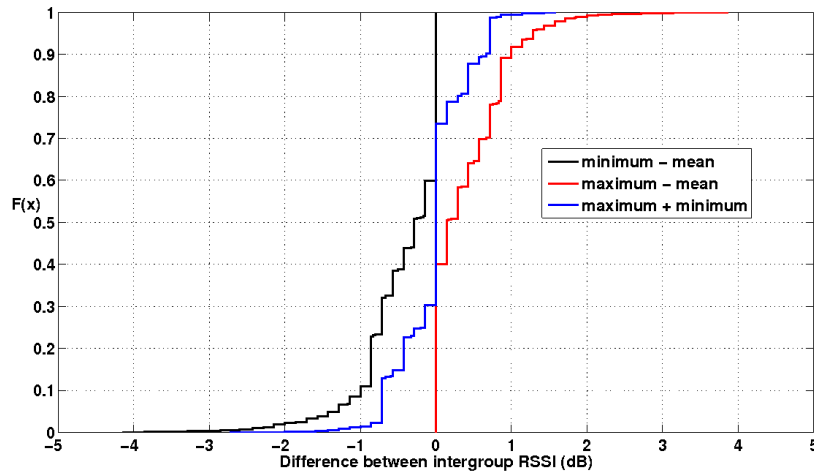


Figure 2.3: Cumulative distribution function plots for the error on the average RSSI.

The black line in figure 2.3 represents the cumulative distribution frequency (cdf)plot of the minimum values minus the corresponding mean for all sending-receiving pairs. It shows that in 40% of the cases the minimum is equal to the mean. When this occurs, all corresponding sending-receiver measurements have the same RSSI-values. The minimum difference equals -4.1 dB. Likewise, the red line represents the cdfplot of the maximum values minus the corresponding mean for all sending-receiving pairs. Again, in 40% of the measurements, there is no difference between the RSSI-measurements in the packet group. The maximum difference is 3.9 dB. The blue line in figure 2.3 represents the cdfplot of the maxi-

mum plus minimum RSSI-values in the inter packet group for all sending-receiver pairs. Almost all values are between plus and minus 1 dB. This illustrates the fact that in our environment a too large inter packet group reading will be accompanied by a too small inter packet group reading. Averaging will further increase the precision to less than ± 1 dB.

Shadowing (occurring when the signal path between the transmitter and receiver is obstructed), is usually modeled as a slowly time varying random process [10]. Averaging now needs to be performed over a larger time. Furthermore, if the data has zero bias, it is interesting to use “truncated” averaging [11]. This is an averaging where the extreme values of the RSSI-measurements (both the smallest and the highest) are ignored. (Please note the median is an example of the truncated mean. Here all, but the most central values are ignored.) Filtering, e.g. with a Kalman filter (see future work in section or an unscented particle filter [12]) efficiently polish up corrupted data. Because we measure an (averaged) accuracy of ± 1 dB in our data, we keep this filtering as future work.

2.3.2 Multipath fading

Multipath fading is the result of radio waves reaching the receiver by two or more signal paths, each with its own time-delay. Depending on the relative phase shift of the waves, these waves either reinforce or undermine each other. This results in a constructive or destructive multipath fading respectively. Causes for multipath fading include refractions and reflections [13].

In order to study the effect of multipath fading in an indoor environment, a test was set up using the w-ilab.t test bed, consisting of nodes 2, 5, 8, 12, 15, 20 and 21 (see the upper part of figure 2.1 (a)). These nodes are collinear and a clear line of sight exists. Every node broadcasts 240 packets. Every packet consists of 100 bytes, the inter packet delay is 25 ms. This brief recovery time between packets allows devices to prepare for reception of the next packet. Transmission is at channel 26 in order to avoid Wi-Fi interference (see further in section 2.3.3). Upon swapping sending nodes, the test bed is idle for 3.5 s. The test is repeated for transmit power levels of 0, -1, -3, -5, -7, -10 and -25 dBm. Every second available data was sent to our log file.

Figure 2.4 shows the RSSI in function of the distance on a semi-logarithmic scale between node 2 and the others at different transmit power levels. The graph also contains the RSSI predicted by the model for a sender transmitting at 0 dBm according to the IEEE 802.15.4 standard [14]. The measured RSSI does not decrease monotonously with distance: e.g. for a transmit power of 0 dBm around a distance of 20 m, a RSSI of -35 dBm is measured and at the shorter distance of about 10 m a RSSI of -47 dBm is observed. Furthermore, at the distance of 20 m the model in the IEEE 802.15.4 expects a RSSI of -70 dBm, giving a difference

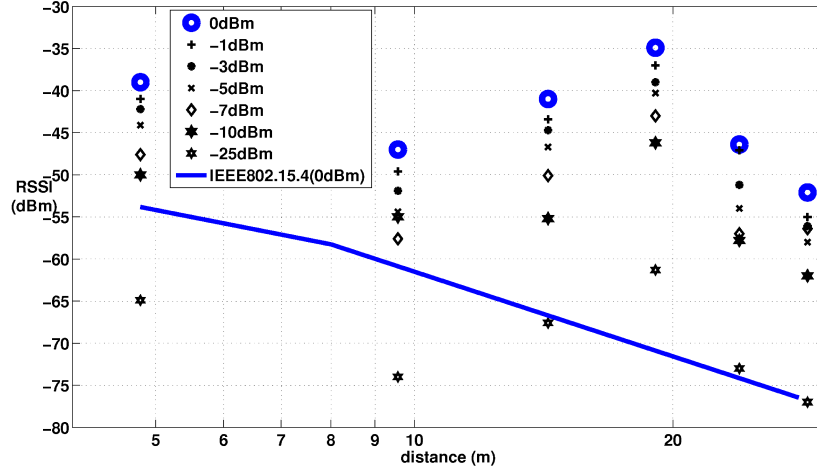


Figure 2.4: Capricious line of sight (LOS) RSSI multipath fading in corridors for node 2 at different power levels.

of 35 dB with the measurement. The measured signal is much too strong, even at small distances. This effect occurs in all corridors and cannot be explained by many theories:

- 1 Walls have attenuation and what we notice is a stronger signal.
- 2 The standard deviation on measurements due to time-fading is only a few dB [15] and what we see is a very large spread.
- 3 The RSSI accuracy in the CC2420 datasheet is much better than the measured spread [16].
- 4 Fluctuations in transmit power are not an issue because figure 2.4 only considers one transmitter (node2).
- 5 The spread on receiver sensitivity of the CC2420 also does not exceed a few dB and is therefore less than the observed spread.

The shape in figure 2.4 with minima and maxima is also found in other practical experiments in corridors [17] and can only be justified with diffraction and the many constructive scatterings and reflections by the conductive metal ceiling and nearby walls in these corridors, acting as “street canyons” [18].

In indoor environment, multipath fading is a very disturbing factor for localization.

2.3.3 Interference

Electromagnetic interference is the disturbance that affects an electrical circuit due to electromagnetic radiation emitted from an external source. Our sensor nodes work in the 2.4 GHz ISM frequency band, as outlined in section 1.4 and discussed in section 2.2. The disturbing sources in this frequency band include:

- Bluetooth (IEEE 802.15.1)
- Wireless USB version 2 (IEEE 802.15.3)
- Wi-Fi (IEEE 802.11)
- microwave ovens
- other sources, like some cordless phones and RF motion detectors

2.3.3.1 Study of interference in the IEEE 802.15.4 standard

The interference problem is recognized by the IEEE, giving rise to a dedicated standard on the interference in the ISM-band [19]. The IEEE 802.15.4 standard [14] provides several mechanisms that enhance coexistence with other wireless devices operating in the 800 MHz, 900 MHz, and 2.4 GHz bands. These mechanisms include

- CCA (Clear Channel Assessment)

The standard [14] provides at least one of the following three CCA methods: energy detection (ED) over a certain threshold, detection of a signal with IEEE 802.15.4 characteristics, or a combination of these methods [20]. The use of the ED option improves coexistence by allowing transmission backoff if the channel is occupied by any device, regardless of the communication protocol it may use.

In [21] coexistence aware clear channel assessment between different standards has also been studied in depth.

- Dynamic channel selection

The same standard provides a “ChannelList” parameter. Out of this list the allowable channels are chosen from. For 2400 MHz band IEEE 802.15.4 networks that are installed in areas known to have high IEEE 802.11b/g activity, the ChannelList parameter can be adjusted in order to enhance the coexistence of the networks.

- Modulation

This standard uses a quasi-orthogonal modulation scheme, where each symbol is represented by one of 16 nearly orthogonal pseudo-random noise sequences. This is a power-efficient modulation method that achieves low signal-to-noise ratio (SNR) and signal-to-interference ratio (SIR) requirements at the expense of a signal bandwidth that is significantly larger than the symbol rate. [10, 14]

Relatively wideband interference would appear like white noise to an IEEE 802.15.4 receiver. The detector performance in this case is similar to noise performance, but the overall SIR requirement is lower because only a fraction of the wideband interferer signal power falls within the IEEE 802.15.4 receiver bandwidth [14].

DSSS is an effective processing tool to deal with interferers whose bandwidth is smaller than the bandwidth of this standard. For example, this processing helps to reduce the impact of an IEEE 802.15.1 interferer [14].

- ED and LQI (Energy Detection and Link Quality Indication)

The LQI is a measurement of the received energy level (signal strength) and the SNR (quality) for each received packet [14, 16]. When energy level and SNR data are combined, they can indicate whether a corrupt packet resulted from low signal strength or from high signal strength plus interference [14].

- Low duty cycle

The IEEE 802.15.4 DSSS is a low power and low data rates standard. A localization algorithm runs on low duty cycles. This will make IEEE 802.15.4 devices less likely to cause interference to other standards.

- Channel alignment

The IEEE 802.11b/g standard [22] recommends non overlapping sets of both European and USA channels. It is possible to use IEEE 802.15.4 DSSS channels that fall in the guard bands between (or above) the three recommended IEEE 802.11b/g channels. While the energy in this guard space will not be zero, it will be lower than the energy within the channels; and operating an IEEE 802.15.4 network on one of these channels will minimize interference between systems [14]. There are 14 IEEE 802.11 channels designated in the 2.4 GHz range spaced 5 MHz apart (with the exception of a 12 MHz spacing between channel 13 and channel 14, the latter channel is only used in Japan). Channel 12 and channel 13 are not used in North-America. As the protocol requires 25 MHz of channel separation, adjacent channels overlap and will interfere with each other. Consequently, using only channels 1, 6, 11 is recommended in the US to avoid interference by the IEEE 802.11 channels [22].

IEEE 802.11b/g Channel	Center frequency MHz	$f_{min} - f_{max}$ MHz	Recommended USA Europe	
1	2412	2401 – 2423	x	x
2	2417	2406 – 2428	-	-
3	2422	2411 – 2433	-	-
4	2427	2416 – 2438	-	-
5	2432	2421 – 2443	-	-
6	2437	2426 – 2448	x	-
7	2442	2431 – 2453	-	x
8	2447	2436 – 2458	-	-
9	2452	2441 – 2463	-	-
10	2457	2446 – 2468	-	-
11	2462	2451 – 2473	x	-
12	2468	2456 – 2478	-	-
13	2472	2461 – 2483	-	x
14	2484	2473 – 2495	-	-

Table 2.1: IEEE 802.11b/g channels

IEEE 802.15.4 Channel	Center frequency MHz	$f_{min} - f_{max}$ MHz	Recommended USA Europe	
11	2405	2404 – 2406	-	-
12	2410	2409 – 2411	-	-
13	2415	2414 – 2416	-	-
14	2420	2419 – 2421	-	-
15	2425	2424 – 2426	x	x
16	2430	2429 – 2431	-	x
17	2435	2434 – 2436	-	-
18	2440	2439 – 2441	-	-
19	2445	2444 – 2446	-	-
20	2450	2449 – 2451	x	-
21	2455	2454 – 2456	-	x
22	2460	2459 – 2461	-	x
23	2465	2464 – 2466	-	-
24	2470	2469 – 2471	-	-
25	2475	2474 – 2476	x	-
26	2480	2479 – 2481	x	-

Table 2.2: IEEE 802.15.4 channels

Figure 2.5 shows that the additional selection of the IEEE 802.15.4 channels 15, 20, 25 and 26 results in a minimum interference system for America.

Likewise, the IEEE 802.11 standard [22] recommends the use of channels 1, 7 and 13 for Europe and IEEE 802.15.4 could benefit from the selection of channels 15, 16, 21 and 22 (see figure 2.6). The effort of the IEEE organization to minimize the interference is shaded by the wild growth of other IEEE 802.11 channel suggestions, like in [23, 24].

- Low transmit power

First, this item is illustrated with spectral measurements of a Ubiwave UW-CM-06 device with both full and reduced output power. Finally, the transmit power is discussed.

1 IEEE 802.15.4 Ubiwave UW-CM-06 spectrum

A Targa 1800 Visionary laptop is equipped with Windows 2003 Server and National Instrument LabVIEW 8. Figure 2.7 illustrates this setup. We use this graphical development environment to control the Rohde & Schwarz Spectrum Analyzer model FSP, accessed through an Agilent 82357A USB to GPIB interface and retrieve the spectrum data in .csv format (the virtual instrument RSFSP_getting_startedApp.vi is adapted to our needs). We conductively connect the antenna output connector of the UW-CM-06 with the spectrum analyzer. Through the JTAG the UW-CM-06 receives the correct software: the external antenna is enabled, the power level is set to -7 dBm (TXCTRL-register = 0xA0EF), the frequency is changed progressively from channel 11 to channel 13. Further the antenna connector is connected conduc-

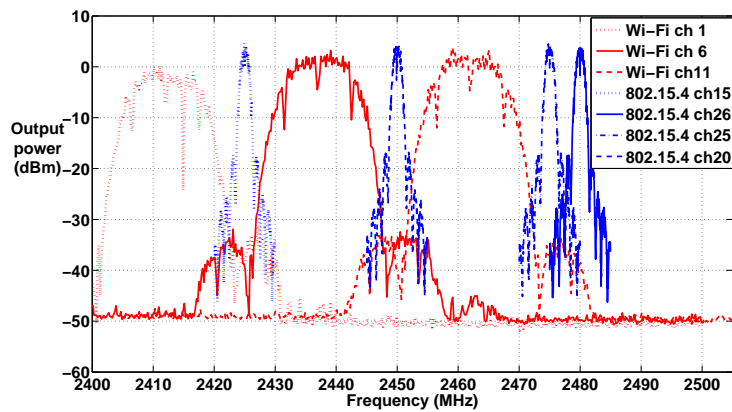


Figure 2.5: USA minimal interference systems consist of Wi-Fi channels 1, 6 and 11, combined with IEEE 802.15.4 channels 15, 20, 25 and 26.

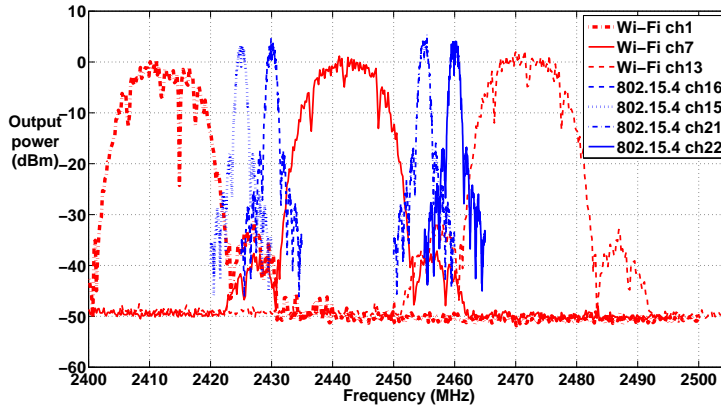


Figure 2.6: European minimal interference systems consist of Wi-Fi channels 1, 7 and 13, combined with IEEE 802.15.4 channels 15, 16, 21 and 22.

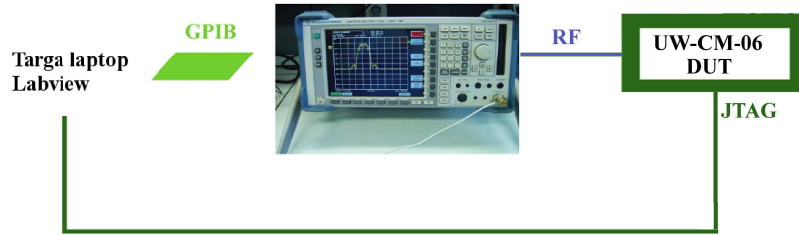


Figure 2.7: Test setup for Ubiwave UW-CM-06 spectrum and power control measurements

tively to the spectrum analyzer and with new GPIB commands the RS-FSP_getting_startedApp.vi is extended to: accepting the Agilent adaptor, a resolution bandwidth of 100 kHz, a viewing bandwidth of 300 kHz, a frequency span of 10 MHz, a RF-level of +10 dBm and reading the spectrum analyzer data in .csv format. The trace of spectrum analyzer is set to MAX HOLD and measurements are captured through the GPIB interface and processed. On figure 2.8 it is shown that IEEE 802.15.4 channel spacing is 5 MHz and that the output power is higher than expected. This is because the internal amplifier is enabled. Also the output power is relatively constant for each channel; it is like the Fourier transform of a raised cosine: the first dip is approximately 1.5 MHz away from the carrier, more dips follow at 1 MHz each (2/3 of this distance), side lobes are at 2, 3, 4, 5 ... MHz away from the carrier. Side lobes are at less than -20, -28, -32, -38 dBc. This device meets the relative 802.15.4 adjacent power specification (< -20 dBc for $|f-f_c| >$

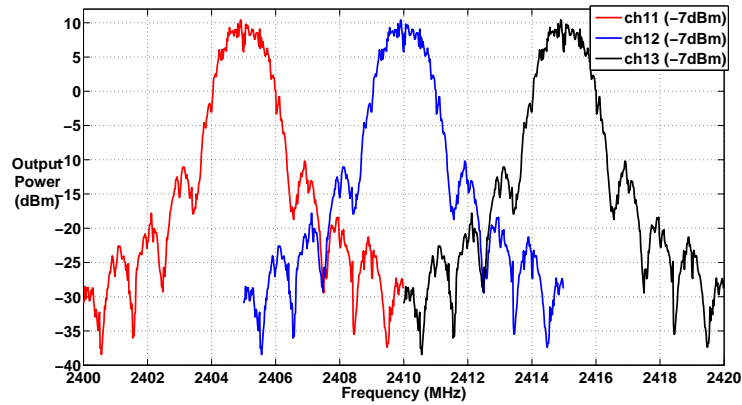


Figure 2.8: IEEE 802.15.4 adjacent channels. Channel spacing is 5 MHz. For these devices, the output level is too high, due to the presence of an internal amplifier.

3.5 MHz) but it does not meet the absolute adjacent power specification (< -30 dBm for $|f-f_c| > 3.5$ MHz) for high output power, because the internal amplifier also shifts the side lobes.

2 IEEE 802.15.4 Ubiwave UW-CM-06 power control.

For channel 12 the output power of the CC2420 chip is respectively set to

- 0 dBm (TXCTRL-register = 0xA0FF)
- -7 dBm (TXCTRL-register = 0xA0EF)
- -15 dBm (TXCTRL-register = 0xA0E7)
- -25 dBm (TXCTRL-register = 0xA0E3)

Figure 2.9 illustrates the measured output power for the different power levels and the corresponding power spectral density (PSD) masks.

- The shape of the curves is very similar for all power levels, except for the highest power. A saturation effect is observed. At this point it is not clear whether the CC2420 or the external amplifier is saturating. Another test on a Silicon Laboratories 2.4 GHz transmitter board [25] without an internal amplifier shows very comparable saturation effects. (Please note this obsolete kit has been replaced by the Silicon Laboratories Ember kit [26].) Therefore, it can be concluded that this saturation is due to the (common) CC2420 radio chip.
- For non-saturation output it is like the Fourier transform of a raised cosine.

- (a) the first dip is approximately 1,5 MHz away from the carrier, more dips follow at 1 MHz each (2/3 of this distance).
 - (b) side lobes are at 2, 3, 4, 5 ... MHz away from the carrier. Side lobes are at less than minus 20, 28, 32, 38 dBc.
 - (c) this device meets the relative 802.15.4 adjacent power specification (< -20 dBc for $|f-f_c| > 3.5$ MHz)
 - (d) this device does not meet the absolute adjacent power specification (< -30 dBm for $|f-f_c| > 3.5$ MHz for high output power, because the internal amplifier also shifts the side lobes.
- When saturating channel broadens: this is a worst case situation
- (a) the passband is flatter
 - (b) the first dip disappears

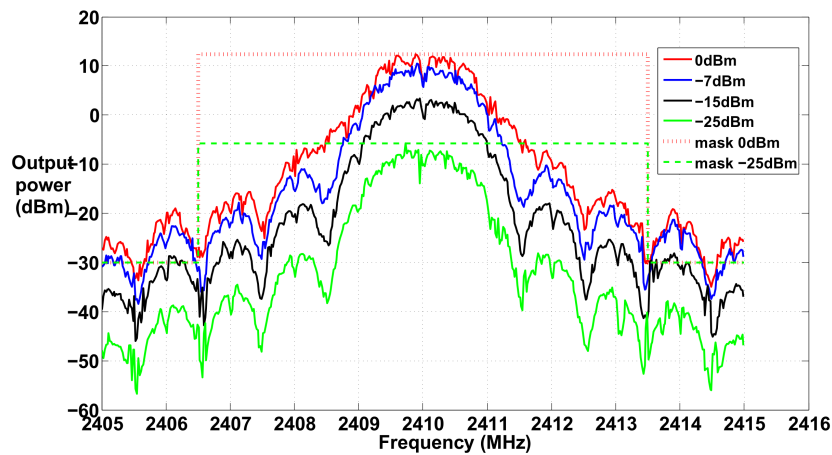


Figure 2.9: IEEE 802.15.4 DSSS power control. When saturating the CC2420 radio chip, the passband widens and some frequency dips disappear.

3 Discussion

The majority of IEEE 802.15.4 devices are expected to operate with transmit powers between -3 dBm and 10 dBm, with 0 dBm being typical. Transmitting at a higher power level results in a higher cost of the low-cost system on chip devices. Furthermore, these higher power levels require expensive filtering in order to meet the regulations for out-of-band emissions. The typical current consumption for a CC2420 chip for an output power of 0 dBm equals 17.4 mA [16]. A very much lower power output of -25 dBm needs a comparable 8.5 mA. At this level, actual transmit power represents a small fraction of the overall power consumed by the transmitter, so there is little benefit in terms

of energy savings to operate below a level of -3 dBm [14]. Please also note that an IEEE 802.15.4 DSSS communication with transmitting power levels of less than -3 dBm is more vulnerable to an IEEE 802.11b/g interferer. The IEEE 802.15.4 DSSS receiver gets a lower signal from the IEEE 802.15.4 transmitter and IEEE 802.11b/g devices transmit at higher power levels than IEEE 802.15.4 transmitters. The CC2420 consumes 18.8 mA [16] in receiving mode. This is slightly above the consumption for transmitting at full power. This further justifies the choice for transmitting at the higher power levels.

Subsection 2.3.3.3 includes a further study of power from an IEEE 802.11b/g interferer received in IEEE 802.15.4 DSSS channels.

- Neighbor piconet capability

A collection of devices occupying a shared physical channel where one of the devices is the piconet master and the remaining devices are connected to it, is called a piconet. Piconets are more associated with Bluetooth [27], but they can also be found in IEEE 802.15.4. This standard offers the possibility to use guaranteed time slots (GTSs). Please note GTSs are situated on the MAC-layer, for a fully treat of the interference, they are also mentioned here. The personal area network coordinator can set aside GTSs specifically for use by other systems. This type of neighbor piconet support capability may further alleviate interference with other systems [14]

2.3.3.2 Frequency spectra of an IEEE 802.11 3-Com wireless Wi-Fi NIC

The quality of localization algorithms increases when physical layer aspects are taken into account. Therefore, this section continues with the study of frequency spectra of a 3-Com wireless Wi-Fi NIC (madwifi drivers already in Kernel) This card is housed in a 600 MHz PC with Scientific Linux (SL04) (CERN, Red Hat based), wireless extensions, and Click-1.5.0 user level. The antenna of the 3-Com NIC is replaced with a connector and conductively connected with the spectrum analyzer. The wireless extensions of the PC are controlled: an IP number is assigned, ad-hoc mode is selected and an ESSID is given. Next, default transmit power (=18 dBm) is set and the frequency is forced to channel 6 (which equals 2437 MHz, see table 2.1). Then a decision on the b or g mode has to be taken: iwpriv mode 2 forces the NIC to IEEE 802.11b and mode 3 is reserved for IEEE 802.11g. Next, the rate is set to 11 Mbps for the IEEE 802.11b and changed from 6 Mbps to 54 Mbps within the IEEE 802.11g. With new GPIB commands the RSFSP_Getting_Started_App.vi is extended to: accepting the Agilent adaptor, a resolution bandwidth of 100 kHz, a viewing bandwidth of 300 kHz, a frequency span 50 MHz, a RF-level of +10 dBm and reading the spectrum analyzer data in

.csv format. The trace of spectrum analyzer is set to MAX HOLD and measurements are captured through the GPIB interface and processed.

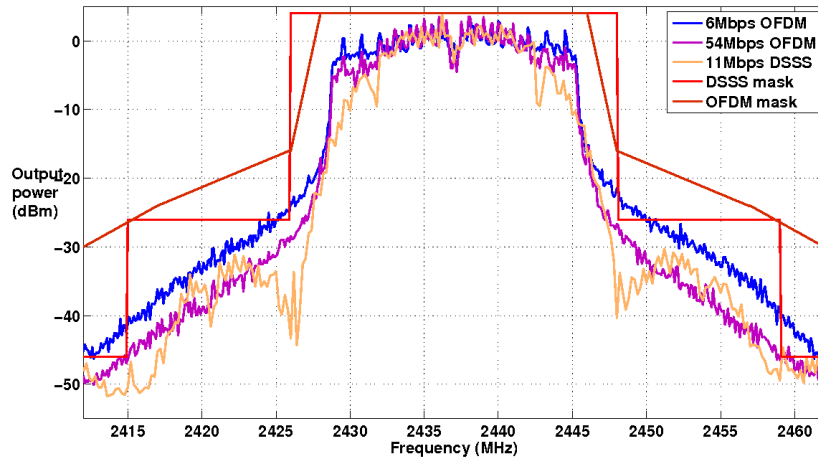


Figure 2.10: The Wi-Fi transmitted power as function of the frequency. All measured spectra fulfill its respective frequency masks.

In figure 2.10 the typical DSSS waveform for IEEE 802.11b modulated at 11 Mbps and the more rectangular spectrum for IEEE 802.11g, modulated with 6 Mbps and 54 Mbps is observed. The two upper curves give the very important transmit power spectral density (PSD) masks for IEEE 802.11g and IEEE 802.11b respectively. Except for distances from the carrier between 9 and 11 MHz, the IEEE 802.11b mask is more restrictive than that of IEEE 802.11g. As discussed in [28] and theoretically confirmed in [29] the widest spectrum is for IEEE 802.11g modulated with 6 Mbps. An important conclusion of figure 2.10 is that all measured spectra fulfill their respective frequency masks.

2.3.3.3 Power measurement of Wi-Fi into IEEE 802.15.4 channels

Power from a single Wi-Fi channel is spread over multiple IEEE 802.15.4 channels and increases its noise floor. With the spectrum analyzer the power in a 2 MHz bandwidth on each IEEE 802.15.4 channel is measured: with a central frequency of 2405, 2410, 2415 ... 2480 MHz. The frequency span is set to 10 MHz, the reference level is 10 dBm, the resolution bandwidth equals 100 kHz and the viewing bandwidth 300 kHz. The NIC is forced to channel 6 and transmits at full power at 11 Mbps (IEEE 802.11b DSSS), and 6 Mbps or 54 Mbps (both IEEE 802.11g OFDM). The results are shown in Table 2.3. In the second column of it, the frequency offset between the central frequency of the IEEE 802.15.4 and the Wi-Fi channel is given. Between round brackets we find the ranking within the same fre-

quency offset: i.e. IEEE 802.15.4 channel 20 at 2450 MHz is 13 MHz away from the Wi-Fi channel 6 and within a 2 MHz bandwidth the least power is for 11 Mbps DSSS, then follows 54 Mbps and finally 6 Mbps OFDM. Due to noise limitations of the measuring equipment, values at high frequency offsets cannot be measured. Please note this ranking could also be previewed on the spectrum of figure 2.10.

2.3.3.4 Minimum distance Wi-Fi transmitter and IEEE 802.15.4 receiver

The IEEE 802.15.4 standard [14] also presents simulated graphs for the minimum separation between an IEEE 802.15.4 (DSSS) receiver and different interferers using another IEEE standard under the simulation conditions of table 2.4. Table 2.5 summarizes these graphs for a Packet Error Rate (PER) of 10%.

The first column displays the channel offset (in MHz) between the IEEE interferer and the IEEE 802.15.4 receiver. Because IEEE 802.15.1 (Bluetooth) uses frequency hopping in the complete ISM-band, no channel separation can be given. The minimum required separation is 19 m. For devices following the other IEEE standards, the minimum required distance increases with decreasing channel offset. The worst result is for the IEEE 802.11b interferer, where a separation of 57 m is needed at a channel offset of 3 MHz. Please note that the minimum of the minimum required distances equals 1.3 m. This implies that interference cannot be avoided, even at the highest channel separation. When two different wireless tech-

IEEE 802.15.4 Channel	F offset MHz	6 Mbps OFDM	54 Mbps OFDM	11 Mbps DSSS
2405	-32	-40.6 (3)	-45.5 (2)	-47.0 (1)
2410	-27	-39.3 (3)	-44.9 (2)	-47.0 (1)
2415	-22	-30.6 (3)	-36.7 (2)	-43.0 (1)
2420	-17	-21.7 (3)	-28.2 (2)	-28.3 (1)
2425	-12	-14.6 (3)	-20.2 (2)	-26.4 (1)
2430	-7	9.9 (3)	6.7 (2)	5.9 (1)
2435	-2	11.8 (2)	11.7 (1)	12.5 (3)
2440	3	12.1 (2)	11.4 (1)	12.7 (3)
2445	8	9.3 (3)	6.4 (2)	1.8 (1)
2450	13	-13.9 (3)	-21.7 (2)	-24.5 (1)
2455	18	-21.0 (3)	-28.8 (1)	-28.4 (2)
2460	23	-30.9 (3)	-37.8 (2)	-41.0 (1)
2465	28	-38.8 (3)	-42.1 (1)	-41.6 (2)
2470	33	-42.0 (3)	-44.9 (2)	-45.0 (1)
2475	38	-	-	-

Table 2.3: Power measurement in 2MHz bandwidth (in dBm) of a Wi-Fi transmitter (channel 6) into IEEE 802.15.4 channels

IEEE 802.15.4 communication link	
channel model	$d = 10^{(Pt - Pr - 40.2)/20}$ for $d < 8m$ $d = 8 \cdot 10^{(Pt - Pr - 58.5)/33}$ for $d > 8m$
receiver sensitivity	-85 dBm
transmit power	0 dBm
transmit mask	see figure 2.9
packets	22 bytes
Bit error rate	$\frac{8}{15} \cdot \frac{1}{16} \cdot \sum_{k=2}^{16} (-1)^k \binom{16}{k} e^{20 \cdot SINR \cdot (\frac{1}{k} - 1)}$
desired signal	10 dB above receiver sensitivity (-75 dBm)
receiver bandwidth	2 MHz
Interferers	
transmit power IEEE 802.15.1	0 dBm
transmit power IEEE 802.15.3	8 dBm
transmit power IEEE 802.11b	+14 dBm
transmit mask IEEE 802.15.1	see section 7.2.3.1 in Standard [27]
transmit mask IEEE 802.15.3	see section 11.5.3 in Standard [30]
transmit mask IEEE 802.11b	see figure 2.10
packets IEEE 802.15.1	1024 bytes
packets IEEE 802.15.3	1024 bytes
packets IEEE 802.11b	1024 bytes

Table 2.4: Overview of simulation conditions

F offset (MHz)	IEEE 802.15.1	IEEE 802.15.3	IEEE 802.11b
NA to IEEE 802.15.1	19		
2		42	
3			57
17		7.8	
22			6.4
27		1.3	
47			2.5

Table 2.5: Simulated minimum required distance in meter between an IEEE 802.15.4 DSSS receiver and different IEEE interferers (PER=0.10)

nologies are housed in the same package, the only solution is a power reduction of the interferer and a kind of channel access control, as discussed in section 2.3.3.1. Two network interface cards that are housed in the same personal computer face the same problem [31]. The upper part of table 2.4 presents the characteristics of the IEEE 802.15.4 communication link, the lower part shows data on the interferer. Please note that there is only one expression for the bit error rate. This is explained

by the fact that there is only one allowed constellation diagram (O-QPSK) in the 2.4 GHz IEEE 802.15.4 DSSS standard. The calculation of this bit error rate follows the approach outlined in section C.3.2 of the IEEE 802.15.2 standard [19]. The higher distance separation for an IEEE 802.11b interferer (compared to an IEEE 802.15.3 interferer) in table 2.5 can be explained by the higher transmit power of the former (see table 2.4). The IEEE 802.15.4 standard doesn't provide data for an IEEE 802.11g interferer. The remainder of this section begins with an empirical study to deal with this problem, and the minimum distance between an IEEE 802.11b/g interferer and an IEEE 802.15.4 receiver is presented. Figure 2.11 shows the test setup.

A 3-Com wireless Wi-Fi NIC (madwifi drivers already in Kernel) is housed in a 600 MHz PC with Scientific Linux (SL04) (CERN, Red Hat based), wireless extensions, Click-1.5.0 user level and a wired NIC with VNC server to control the wireless card out of the Qosmotec box. These perfectly shielded metal boxes are used to completely annihilate the influence of interferences, meanwhile giving the opportunity to optically connect to the equipment placed inside. An IEEE 802.15.4 connection with two TMote Sky modules housed in other Qosmotec boxes will be set up. A TMote connect gateway interfaces the TinyOS 2.0 TMote Sky IEEE 802.15.4 modules to Ethernet.

A Targa 1800 Visionary laptop is equipped with Windows 2003 Server, a wired NIC with a VNC client and National Instrument LabVIEW 8. We use this graphical development environment to control the Rohde & Schwarz Spectrum analyzer model FSP, accessed through an Agilent 82357A USB to GPIB interface and take back the spectrum data in .csv format (RSFSP_getting_startedApp.vi is adapted to our needs). We supervise the emissions with a D-Link wireless NIC and Wireshark Airopeek sniffer. To exclude unwanted radiation the devices under test are sealed within a shielded Qosmotec box. The laptop also holds Netbeans IDE 6.0 and the Radioperf tool [32]. This tool is a Java interface application, where the designers made it possible to remotely control and interrogate the TMote Sky modules. First we will use Radioperf to set the number of bytes per packets and number of packets per second of the IEEE 802.15.4 transmitter and then Radioperf reads the total received and lost packets. Finally the laptop will also be used to remotely connect to the Qosmotec Air Interface Simulator (AIS). This simulation tool for lab usage does away with imprecise manual attenuators and allows for controlled, repeatable, yet easy-to-use simulation [33]. Here, the Qosmotec system is used to control the attenuation between the devices under test and control the distances in absence of interference.

The experiment continues with the setup of the system of figure 2.12 using Qosmotec splitters and attenuators. All devices under test are put in a separate Qosmotec box. The power of the IEEE 802.15.4 transmitter is split and conducted to two attenuators. The first attenuation path leads to the IEEE 802.15.4 receivers

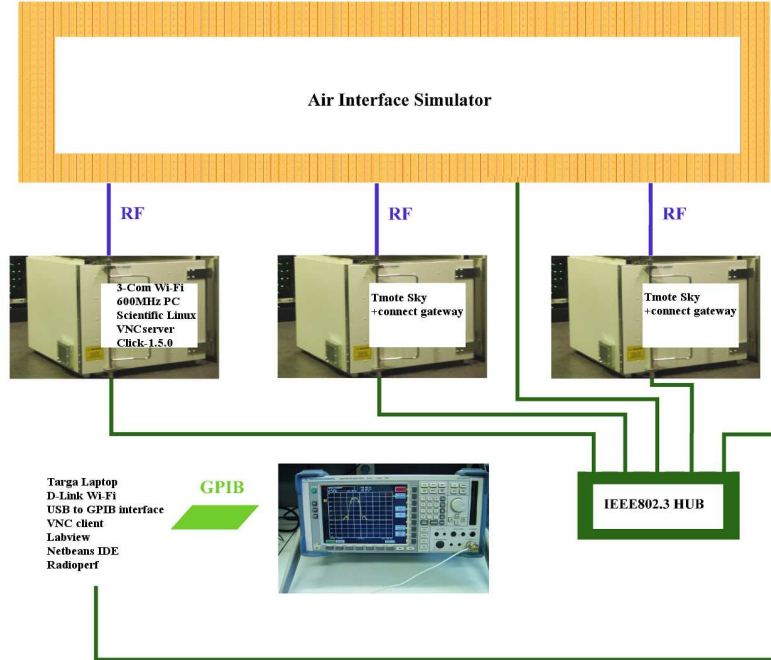


Figure 2.11: Test setup with shielded QOSMOTEC boxes for testing interference of a Wi-Fi interferer on an IEEE 802.15.4 link.

with an Air Interface Simulator fixed attenuation of 68 dB. The second attenuation path leads to the splitter/combiner of the Wi-Fi transmitter. Its AIS attenuation is set to the maximum (92 dB). At this level both transmitters will not cease transmitting packets. The other leg of the splitted Wi-Fi power is sent to an AIS variable attenuator, representing the variable distance between Wi-Fi interferer and IEEE 802.15.4 receiver. The IEEE 802.15.4 receiver combiner adds the attenuated signals from the IEEE 802.15.4 and Wi-Fi transmitters.

Table 2.6 gives an overview of how the measurement conditions relate to the simulation conditions of the IEEE 802.15.4 standard [14]. Both measurement and simulation use the same channel model. Our measurement conditions are more stringent than the simulations in the standard: our receiver is more sensitive and in analogy with analog receiver measurements we only went 3 dB above the sensitivity level. Furthermore, a more powerful interferer produces higher side lobes and more out of band noise. It is also easily verified that with the given packet sizes and packet repetition rates an IEEE 802.11 packet with sufficient power cannot find a time slot between two IEEE 802.15.4 packets and will therefore always cor-

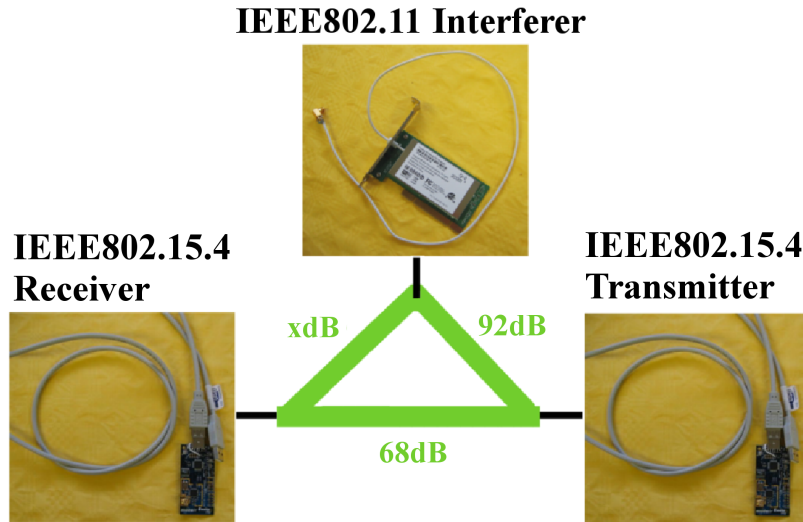


Figure 2.12: Equivalent test setup for determining the minimum required distance separation.

rupt an IEEE 802.15.4 packet. The PER is larger for bigger packet sizes because the larger packets are more prone to errors than the smaller ones [34]. As already mentioned, the standard does not provide data for an IEEE 802.11g interferer.

The wireless extensions force the NIC at full power to channel 1 (2412 MHz) and change the modulation rate from 11 Mbps, 6 Mbps and 54 Mbps. The Click software creates 4000 packets per second with a length of 100 bytes each. With the help of the earlier described Radioperf tool an IEEE 802.15.4 communication is set up at frequency offsets of 3, 8, 13, 18, 48 and 68 MHz away from the Wi-Fi transmitter. Each 30 ms the IEEE 802.15.4 transmitter sends a packet of 100 bytes and the receiver logs the number of received valid IEEE 802.15.4 packets in .csv format. For each modulation rate and each frequency offset the attenuation is changed and its corresponding number of received packets is logged. Obtaining one attenuation-received packets pair takes a 5 minutes measurement of 10000 sent IEEE 802.15.4 packets.

Basically the equivalent of the measurements is given in figure 2.13. All other equipment between the IEEE 802.15.4 transmitter and receiver gives an additional and fixed attenuation of 21 dB and with the 68 dB of the attenuator, this results in a total attenuation of 89 dB. Using the channel model of Table 2.6 we find a distance of 75 m. This model is also the basis to correlate the minimum required separation of the Wi-Fi transmitter and the IEEE 802.15.4 receiver out of the difference between the AIS attenuations.

	measurement conditions	simulation in IEEE 802.15.4 standard
channel model		
for $d < 8m$	$d = 10^{(P_t - P_r - 40.2)/20}$	idem
for $d > 8m$	$d = 8 \cdot 10^{(P_t - P_r - 58.5)/33}$	idem
receiver sensitivity	-92 dBm	-85 dBm
desired signal		
absolute	-89 dBm	-75 dBm
above receiver sensitivity	3 dB	10 dB
transmit power		
IEEE 802.11g	+20.1 dBm (6 Mbps)	n/a
IEEE 802.11b	+20.8 dBm	+14 dBm
IEEE 802.15.4	0 dBm	0 dBm
receiver bandwidth	not measured	2 MHz
transmit mask		
IEEE 802.11b	see figure 2.10	see figure 2.10
IEEE 802.11g	see figure 2.10	n/a
packets IEEE 802.11b	100 bytes, 4000packets/s	1024 bytes
packets IEEE 802.11g	100 bytes, 4000packets/s	n/a
packets IEEE 802.15.4	100 bytes, 30packets/s	22 bytes

Table 2.6: Overview of measurement conditions

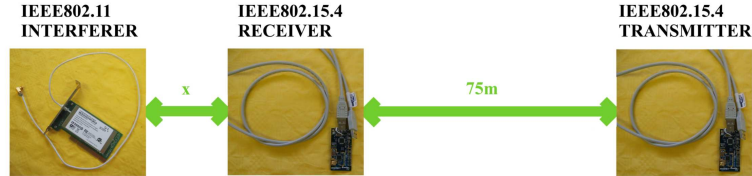


Figure 2.13: Equivalent distance measurement setup.

The results of the post processing is given in figure 2.14 for frequency offsets of 3, 18 and 48 MHz and figure 2.15 for 8, 13 and 68 MHz.

Table 2.7 gives the minimum required distance in meter between a Wi-Fi interferer and an IEEE 802.15.4 receiver at a PER of 0.10 for modulation rates of 6, 54 and 11 Mbps under the measurement conditions of table 2.6. In the first column we find the frequency offset in MHz. At a frequency offset of 18 MHz i.e. a 6 Mbps (OFDM) modulated Wi-Fi needs to be more than 26 m away from the IEEE 802.15.4 receiver in order to obtain a PER of maximum 0.10. When changing

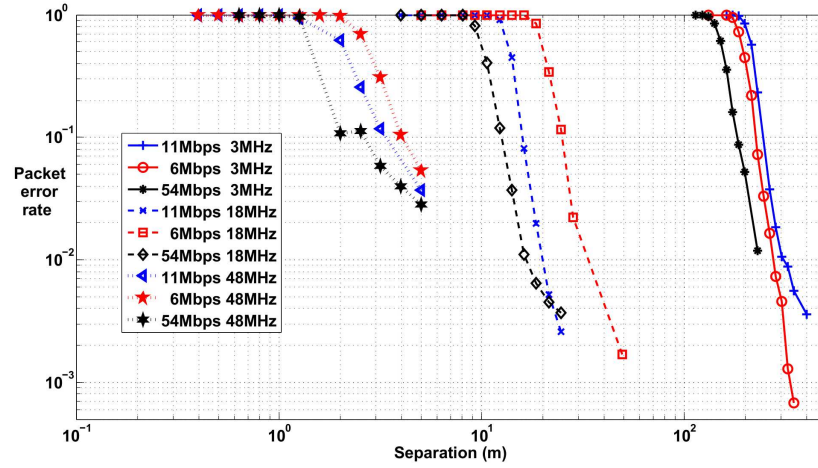


Figure 2.14: Worst case distance required between LAN interferer and IEEE 802.15.4 receiver for frequency offsets of 3, 18 and 48 MHz.

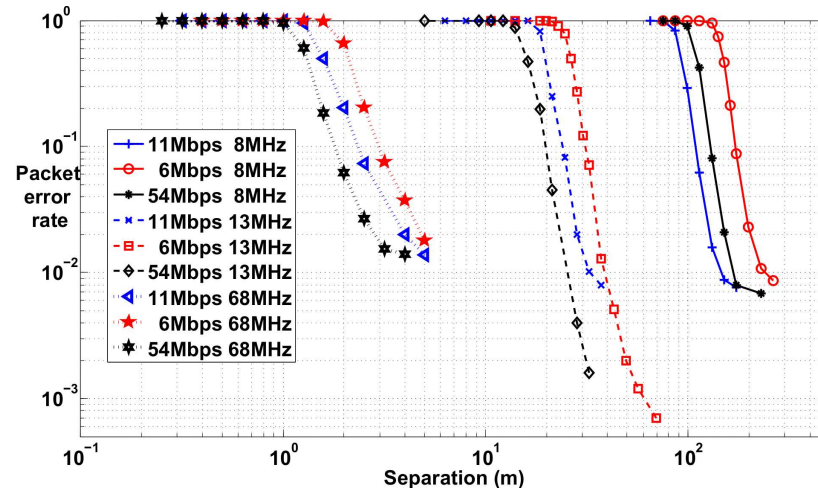


Figure 2.15: Worst case distance required between LAN interferer and IEEE 802.15.4 receiver for frequency offsets of 8, 13 and 68 MHz.

the modulation to 54 Mbps (OFDM) the interferer can come closer (13 m) for the same PER. Please note the 11 Mbps (DSSS) (IEEE 802.11b) performs worse than the 54 Mbps (OFDM) (IEEE 802.11g). This is in accordance to [35]. Furthermore, increasing the frequency offset will have a positive effect on the immunity to interference. For the same channel separation, the values of the last column of table 2.7 are higher than those of the last column of table 2.5. This can be explained

by the fact that the measuring conditions are more stringent than the simulation conditions.

F offset (MHz)	6 Mbps OFDM IEEE 802.11g	54 Mbps OFDM IEEE 802.11g	11 Mbps DSSS IEEE 802.11b
3	240	180	220
8	190	140	105
13	30	20	23
18	26	13	16
38	4	2	3
43	3	1.9	2.3

Table 2.7: Minimum required distance in meter between a Wi-Fi interferer and an IEEE 802.15.4 receiver (PER=0.10)

2.3.3.5 Interference from microwave oven

Figure 2.16 shows the measured frequency spectrum of a 800W Primo-MG1-B microwave oven when boiling two cups of water. The trace of the Rohde & Schwarz FSP spectrum analyzer is set to MAX HOLD. The right-hand side of the figure shows narrowband signals, coming from the free running magnetron. At the left-hand side of the figure these peaks are much closer to each other. This can be explained by the fact that the sizes of the magnetron cavities determine the resonant frequency, and thereby the frequency of emitted microwaves. The frequency is not precisely controllable. The operating frequency varies with changes in load impedance [36], with changes in the supply current, and with the temperature of the tube. Therefore, a very wide disturbing frequency band is encountered, almost filling the complete 2.4 GHz ISM band [37]. The conventional power supply of a microwave oven, like the Primo-MG1-B, consists of a step-up transformer and a diode rectified voltage doubler [38]. This kind of microwave only radiates on the positive peaks of mains supply [39].

In order to study the effect of microwave interference, a test was set up using the RadioPerf tool (see section 2.2). Two IEEE 802.15.4 nodes are put on a table at a distance of 1.4 m and a communication was set up. One node continuously transmits packets of 100 bytes at full power, the inter packet delay is 30 ms. The receiver was put a few centimeters aside the door of the microwave oven. Transmission is at channel 22 (2460 MHz) in order to capture much microwave interference (see figure 2.16). During the experiment, the microwave was switched off and on. Every second available data was sent to our log file.

In figure 2.17 a screenshot of RadioPerf is given: the upper and the lower graph

present the received and the lost packets, respectively.

Up to a time of -255.100 ms there is very little packet loss. Then, switching on the microwave results in a decreased number of received packets and an increased number of lost packets. However, there is no situation where all packets are lost. This can be explained by the facts that the microwave actually uses a kind of frequency hopping and that there is no emission on negative peaks of mains supply.

An example where Clear Channel Assessment is used to improve the resilience to microwave oven interference is given in [40]. Furthermore, when all microwave ovens are fed from the same single-phase electrical power lines (which is usually the case), one could detect the phase of that power supply and design a wireless sensor network that avoids sending packets on the peak of the mains supply.

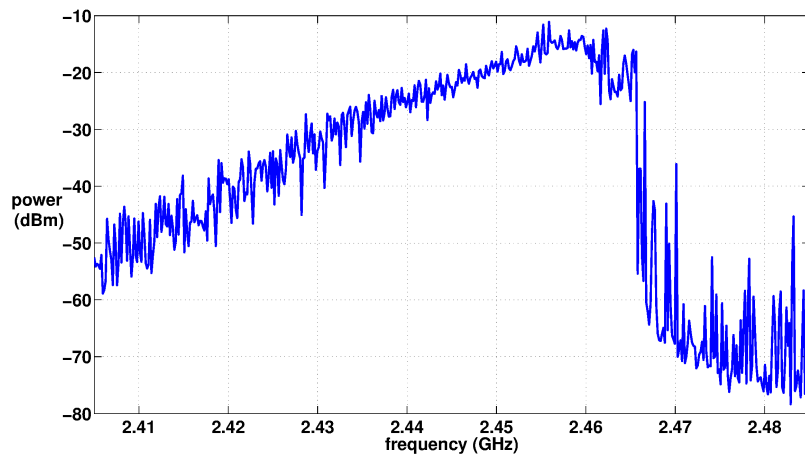


Figure 2.16: The spectrum of a microwave oven almost spans the whole ISM 2.4 GHz frequency band and reaches maximum output around 2.46 GHz.

2.3.4 Radiation pattern

This section discusses the radiation pattern of the w-ilab.t T-mote Sky modules. These modules embody a microstrip planar inverted-F antenna (PIFA). This name refers to the ground faced F-shape of the antenna (see figure 2.18). The feed of the antenna accesses two open transmission lines (TML). The inductance of the first TML resonates with its capacitance to the ground, like in an open dipole. The other open transmission line, TML2, is capacitive loaded at the top with a shorted stub, also forming a resonant structure. This capacitive loading reduces the size of the antenna, making it suitable for mounting on portable equipment. More about the design of a 2.4 GHz PIFA can be found in [41]. The PIFA is currently being used in many radiotelephone handsets, because it is small and has a low profile [42, 43].

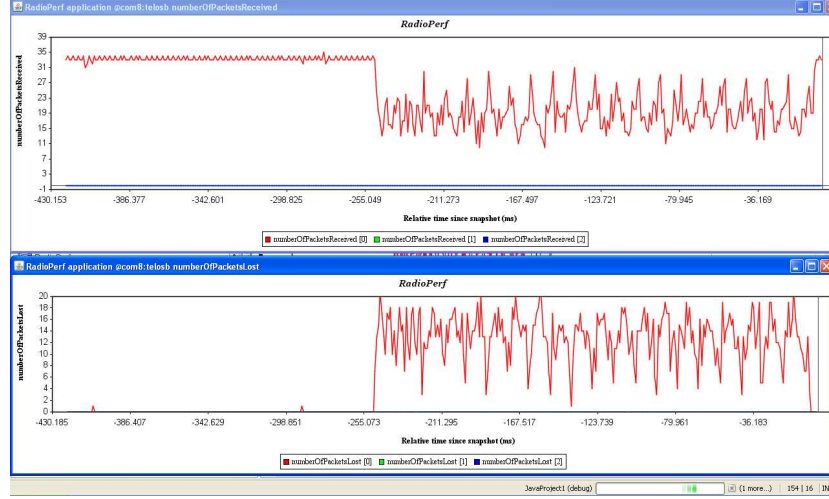


Figure 2.17: A microwave oven with a conventional power supply does not completely block an IEEE 802.15.4 link, because the microwave only causes interference on the positive peaks of the mains supply.

Furthermore, it deals with the shortfalls of the $\frac{\lambda}{4}$ monopole antenna in mobile communication applications and the antenna can reduce the possible electromagnetic energy absorption by the mobile handset users head, because of relatively smaller backward radiation toward the user [44].

Figure 2.19 represents the radiation pattern in the horizontal plane, retrieved from the TMote Sky datasheet [45]. The corresponding position of the TMote is printed on top of the radiation pattern. Please note that the TMote module is presented with a top view, like in [46]. In our w-ilab.t test bed, the nodes are mounted with the top of the printed circuit board down, and the pattern needs to be flipped horizontally. The pattern is quite omnidirectional, except for two dips. The first is at 180 degrees and hence in the direction of the USB connector and antenna-feed (see figure 2.18). This dip is about 24 dB deep, but very narrow. It is also encountered in measurements of a freestanding PIFA, like in [41, 47], therefore, it is attributed to the antenna feed. The second and less deep dip comes at approximately 260 degrees (or 80 degrees if mounted as in our w-ilab.t test bed). Some authors define a Degree of Irregularity (DOI) as the maximum range variation per unit degree change in the direction of radio propagation to model the radiation of the antenna [48]. With this DOI the emitted directional power varies between two values. Others refine this model and present Radio Irregularity Model (RIM), based on parametric data fitting of a Weibull distribution. A method that is more direct, however, measures the antenna pattern in an anechoic chamber and uses this pattern in a localization algorithm, like in [46].

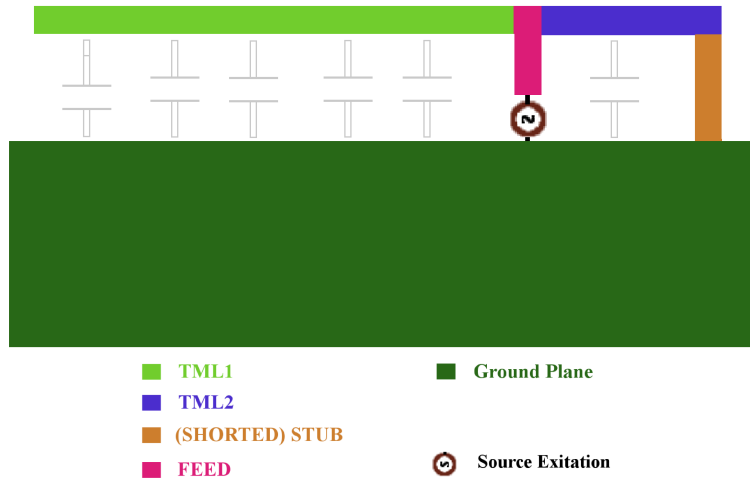


Figure 2.18: Principle of a F-shape antenna: it consists of one open and one stubbed transmission line. The parasitic capacitances to the ground plane are also shown.

2.4 Design of a WSN transmitter

2.4.1 Introduction

Our test bed uses cheap and small TMote Sky modules, as already explained in section 2.2. These modules are adequate and fairly convenient for existing localization algorithms. A major drawback, however, is that they are not designed for a profound study on the physical (and MAC) layer of the OSI model. The CC2420, i.e. directly outputs the modulated RF signal. The inphase signal (I-signal) and quadrature signal (Q-signal) of the O-QPSK modulator are only available in the interior of the chip. This work aims at improving localization algorithms by a better understanding of the physical layer of the IEEE 802.15.4 DSSS standard [14]. How can we comprehend a standard better than by actually building a transmitter that meets the requirements of this standard? This section therefore focuses on an implementation of the IEEE 802.15.4 DSSS standard and describes the design of hardware and software for a WSN transmitter. To be most flexible, the principles of Software Defined Radio (SDR) are used here. SDR is not only a good answer, but also a versatile and efficient solution for future upcoming standards. In SDR, some hardware components are replaced by software, such as mixers, filters, modulators and amplifiers. Other components are impossible to be realized in software, for example some parts of the physical layer, and are therefore implemented in hardware. Existing SDR, like the GNU open source radio project [49], is versatile, but the available hardware is expensive. We decided to build our own

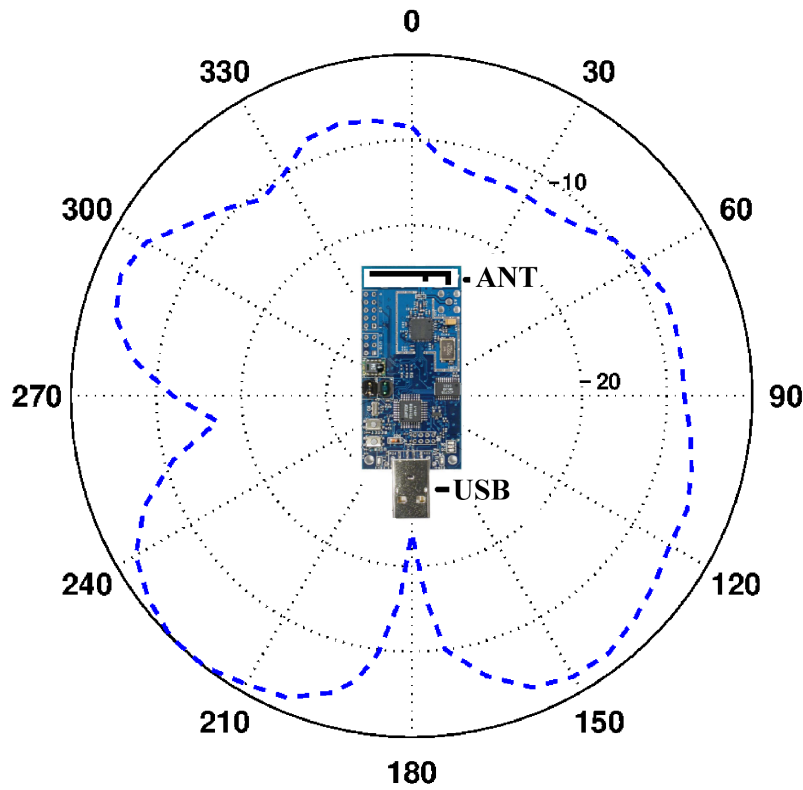


Figure 2.19: Radiation pattern of the T-mote Sky module (top view) with horizontal mounting [45].

hardware and hence also own software. Necessary design tips are described. Here, we build the transmitter pushing both the software and the hardware to their limits. This section is organized as follows. A first subsection will describe the hardware used, whereas the following subsection will give indications for the implementation of the IEEE 802.15.4 protocol in software. In a next subsection, results will be described showing that our design meets the IEEE 802.15.4 DSSS physical layer specifications. Finally, in the last subsection conclusions will be drawn. More about this design can be found in [50, 51].

2.4.2 Hardware

The hardware is based on three main parts: a standard PC, a data acquisition (DAQ) card and a self-made up-converter printed circuit board (PCB). Figure 2.20 gives a schematic view. The PC (with LabVIEW [52] as programming language)

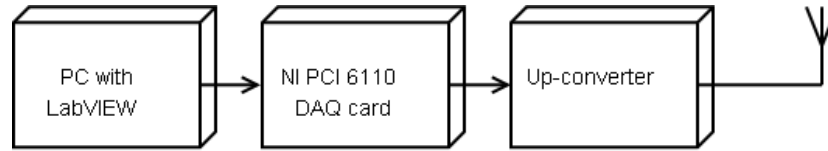


Figure 2.20: The block scheme of the hardware used.

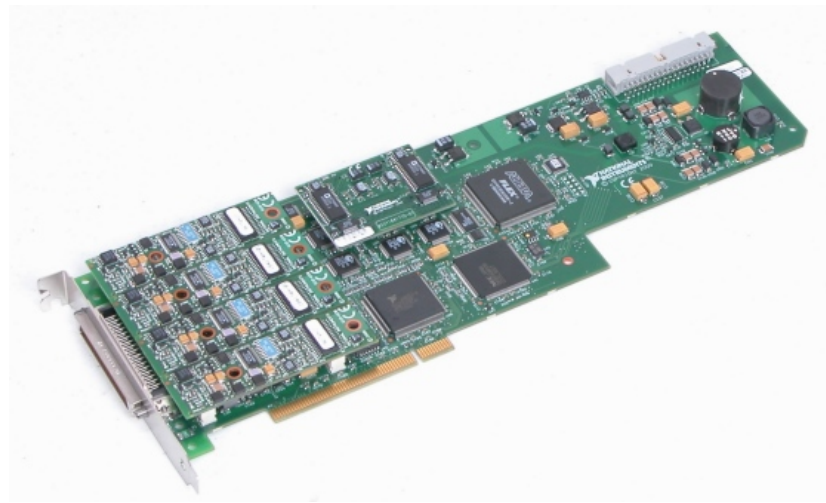


Figure 2.21: The DAQmx card contains both analog and digital inputs and outputs.

generates baseband signals, which are converted via the PCI-bus of the PC through the DAQ card into analog signals. In this work, a National Instruments PCI 6110 (or DAQmx card, shown in figure 2.21 [53]) has been chosen, which contains both analog and digital inputs and outputs. If two analog outputs were used, a sample rate of 2.5 MS/s could be achieved. The up-converter is used to shift these signals to the 2.45 GHz ISM frequency band, tunable by the software and generated with a classical digital Phased Locked Loop (PLL) (which will be explained later). The PCB also contains a quadrature modulator, responsible for the mixing of the I- and Q-signal and for the frequency shift to an RF signal. In this design, an AD8349 [54] is used. This high performance quadrature modulator is used as a single-stage up-converter with a high output power and very low noise floor. Because the modulator requires differential signals, whereas the output of the DAQ card is single ended, two additional differential amplifiers are integrated. A picture of the design can be found in figure 2.22, where most of the connectors are used for debugging purposes.

Like already said, the PCB also contains a classical digital PLL, which is programmed via the digital outputs of the DAQ card. This Phase Locked Loop is a

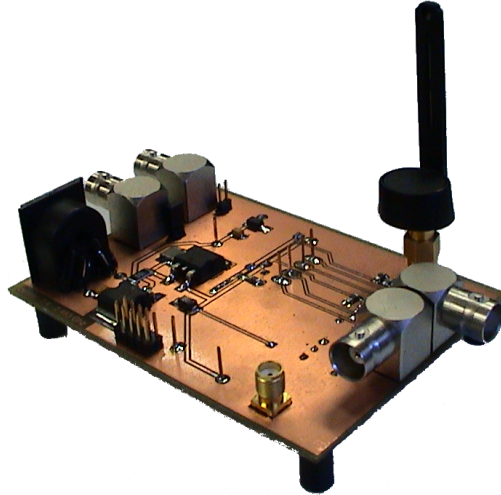


Figure 2.22: The transmitting PCB contains the up-converter, and can easily be debugged.

tuning system with a closed loop that generates an output signal as function of the frequency and the phase of the input signal. Usually, this is done automatically by changing the frequency of a Voltage Controlled Oscillator (VCO), resulting in a signal with the same frequency and phase as the input signal [55]. In the follow-

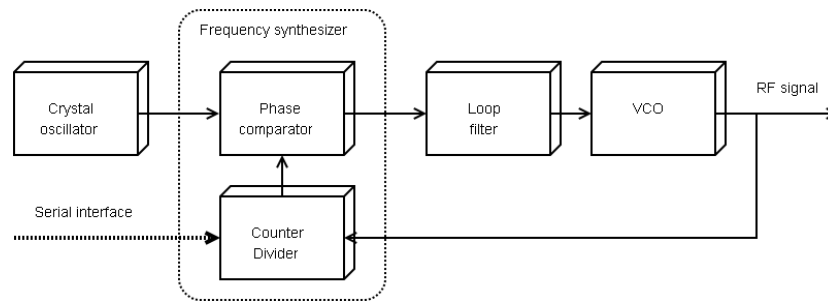


Figure 2.23: The block scheme of the classical digital PLL.

ing paragraphs the design and the operation of the PLL are described, following the design rules suggested in [56]. As can be seen in figure 2.23, the input signal is a clock signal which is a crystal oscillator in this design. This input signal will be compared with the divided output signal in the phase comparator (which is part of the frequency synthesizer). In order to generate high frequencies, this division is necessary. When the phase and/or the frequency of both signals differs, the frequency synthesizer generates an output signal. It will be higher if the fre-

quency/phase is too small and lower if frequency/phase is too large. In order to reduce the jitter coming from glitches of the charge pump in the frequency synthesizer, this signal is also filtered with a low pass filter. The high frequencies of the glitches are hence cut off with the filter. The input signal changes pass rather fast through a filter with a large bandwidth, resulting in an unstable PLL. A trade off between a fast and a stable PLL should be made. For our application, a fast PLL is not necessary (it is only used to change the carrier frequency), and hence a second order filter with a limited bandwidth of 144 Hz is selected. After filtering this signal in the loop filter, the Voltage Controlled Oscillator (VCO) gives the output signal. The chosen VCO is the CVCO55BE-2400-2670 [57], where a frequency of 2.3 GHz to 2.67 GHz can be obtained using a tuning voltage of 0 V to 15 V. In fact, ADF4113HV [58] is the frequency synthesizer used here. It is a chip with several programmable counters, a phase detector and a charge pump. The counters are programmed by the integrated serial interface, via the LabVIEW software and the digital outputs of the DAQ card. This solution gives us the possibility to highly accurately generate a frequency in a wide range. Therefore, the frequency synthesizer requires an as stable as possible oscillation frequency of 5 MHz to 150 MHz. Hence, a standard crystal oscillator of 10 MHz is selected. An external reference signal is also possible, but this is a less flexible solution. The final electronic de-

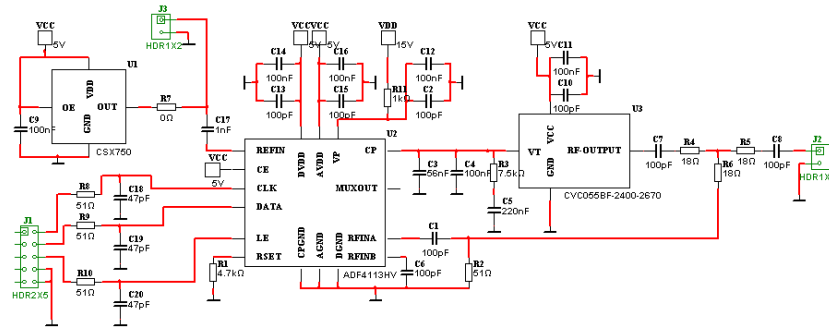


Figure 2.24: The electronic design of the PLL including frequency synthesizer, loop filter and VCO.

sign of the PLL with all necessary peripheral components is shown in figure 2.24. This is also realized in a part of the transmitting PCB in figure 2.22.

2.4.3 Software

Also in the software some additional design steps should be taken, which are described in this paragraph. As already said in the introduction, we are focusing here on the IEEE 802.15.4 standard in the 2.45 GHz frequency band. The binary information to be transmitted is in a first step translated into different symbols (as

indicated in figure 2.25) and in a next step into chips (chip length 32) with predefined pseudo-random noise sequences. The standard also describes for every chip bit a half-sine pulse shaping, in order to reduce the effective bandwidth. A side effect is the additional side lobes, which need to be filtered in the final signal in order to lower the power in the other frequency bands. Finally, the modulation scheme is O-QPSK (Offset Quadrature Phase Shift Keying), where the offset is necessary to reduce the amplitude fluctuations between the 4 different phases and hence 4 different symbols. The offset is realized by an additional time shift of a half symbol length between the I- and Q-signal, resulting in the fact that both signals cannot change in the same time slot and hence reduces the maximal phase transitions to 90. All those steps are implemented in LabVIEW and fed into the DAQ card. The outcome of the software is two signals (I and Q) and serves as input for the transmitting PCB. In order to validate the already taken design steps, the

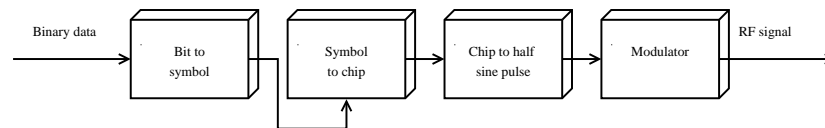


Figure 2.25: The block scheme for an IEEE 802.15.4 signal.

software is written to generate a simple ACK (Acknowledgment) frame. Figure 2.26 illustrates the schematic overview of the acknowledge frame [14, 50]. It con-

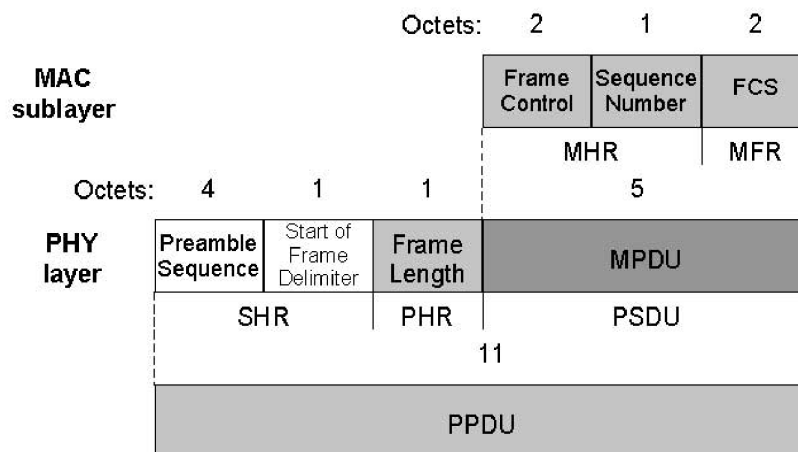


Figure 2.26: Schematic view of the acknowledge frame.

tains 3 main parts: a SHR (synchronization header), a PHR (physical layer header) and a PSDU (physical layer service data unit). The SHR is used for synchronization and contains a preamble sequence of 32 binary zeros and a fixed sequence as

start-of-frame delimiter 1100101. The PHR gives the frame length in bytes of the following PSDU, which is limited to 127 bytes (resulting in a maximal length of 7 bits). An ACK frame always contains 5 bytes, hence the PHR is 0100000. For an ACK frame, the PSDU only contains a header and a footer. This header starts with 3 bits frame type (which is 010 for an ACK frame), followed by the security bit (for an ACK frame always 0) and a frame pending bit (no additional frames are expected for an ACK frame, hence 0). Then there is a bit reserved for an ACK frame request, which is for an ACK frame of course equal to 0. The intra PAN ID indicates that the frame remains in the same network, which is always the case for an ACK frame and hence results in 1. Normally, this is followed by the address fields, but for an ACK frame no addresses are included. Instead, a sequence number is added. In our application a random sequence number is selected. The PSDU footer is a frame check sequence (FCS) based on a 16 bit CRC (Cyclic Redundancy Check).

2.4.4 Results

In figure 2.27, both the I- and Q-signal are plotted as function of the time, before any pulse shaping. This signal is a software generated ACK frame like described above and measured at the output of the DAQ card. From both I- and Q-signal, the according spectral behavior can be generated with a spectrum analyzer. There is at this stage no up-converter involved, resulting in a baseband signal.

The signals of figure 2.28 are used as inputs of the transmitting PCB. This up-converter mixes and modulates both signals. The spectral output is given in figure 2.29, which shows that the bandwidth is exactly 2 MHz, meeting the prescriptions of the standard. Please note that in these figures, the side lobes of the signal have not yet been filtered. In order to check the validity of a transmitted data packet, we are using a commercially available packet sniffer. The hardware is the CC2420DK Development Kit from Chipcon [59], while the packet sniffer software is freely available. In figure 2.30, transmitted acknowledgment (ACK) frames, like described above, are sniffed. For measurement purposes, different ACK frames are repeated. Please note that the frame check has not yet been implemented correctly, resulting in an error. The packet sniffer software also gives the RSSI (Received Signal Strength Indicator), as an indication of the received power. The distance in the lab measurements was limited to approximately 2 m, resulting in rather large and fixed values for RSSI. In order not to disturb the neighboring channels, two constraints for the transmitted power spectral density should be considered. One constraint (relative limit) is that the power of frequencies differing more than 3.5 MHz should be 20 dB lower than the carrier frequency and the other constraint (absolute limit) should be lower than -30 dBm.

According to figure 2.29, both the absolute and the relative limit are fulfilled.

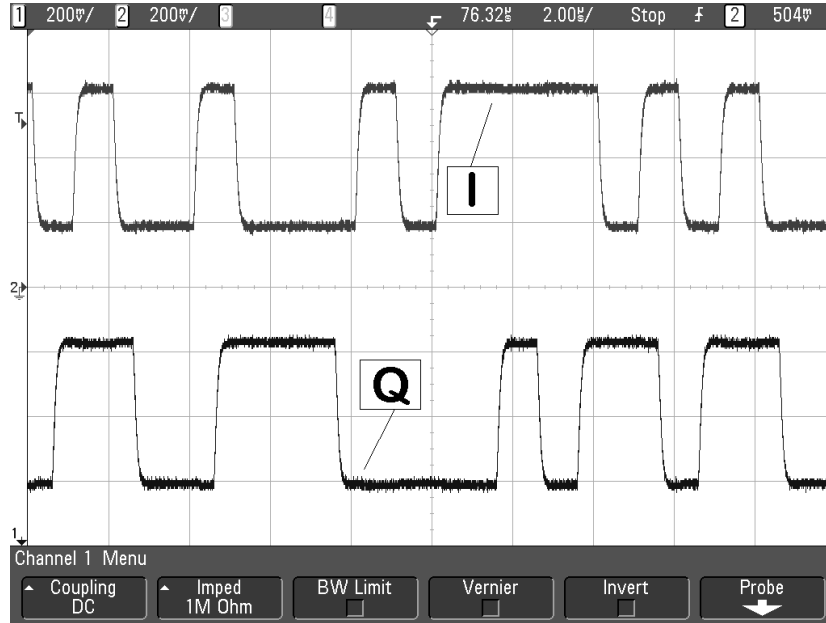
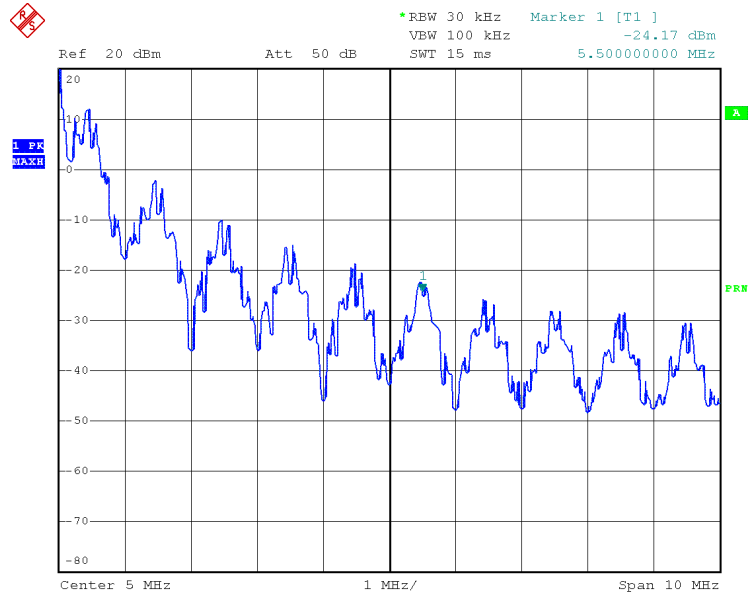


Figure 2.27: I-signal and Q-signal as function of the time.

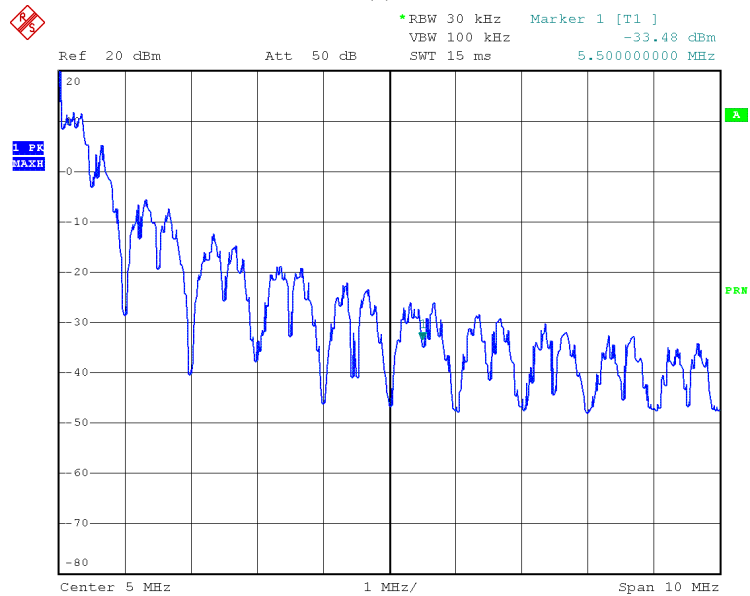
This means that no power is lost in frequencies outside the considered band and that our design is according to the standard.

2.4.5 Conclusion

Based on a DAQmx card (for the necessary I- and Q-signal) and a transmitting PCB, which mixes and modulates the signals, an RF signal is generated. We also proved with time and frequency domain figures and with real transmitted acknowledgment frames that our design meets the IEEE 802.15.4 DSSS physical layer specifications. This section helped to understand the IEEE 802.15.4 standard better. Indeed, our versatile software based approach improved access to the physical layer. It is clear now that the good spectral response is related to the chirp to half-sine pulse shaping step (prior to O-QPSK modulation). Furthermore, using SDR techniques, the self-made transmitter communicated with commercially available devices. A design of a software defined receiver, based on the same design, would allow to explore the standard more profoundly.



(a)



(b)

Figure 2.28: Spectrum of the I-signal (a) and Q-signal (b). These signals are input for the AD8349 Quadrature modulator.



Figure 2.29: The spectral output of the generated IEEE 802.15.4 signal fulfills the absolute and relative PSD requirements.

2.5 Conclusions

This chapter started with the choice of a suitable ranging technique and wireless technology based on the design goals of section 1.3. The IEEE 802.15.4 DSSS ranging technique has been selected, because it has a wide extended range. Furthermore, it is designed for a simple, low-cost communication network that allows wireless connectivity in applications with limited power and relaxed throughput requirements. The main objectives of this standard are ease of installation, reliable data transfer, short-range operation, extremely low cost, and a reasonable battery life, while maintaining a simple and flexible protocol. Having 64-bit extended addresses, the standard scales extremely well. Other UWB based systems require complex timing techniques, and might perform better than IEEE 802.15.4 DSSS systems on accuracy and robustness. In our design of a localization algorithm (see chapter 4 and chapter 5) these characteristics will be given extra attention. The IEEE 802.15.4 DSSS ranging technique definitely outperforms the other techniques in the other design goals.

In a next section, the w-ilab.t test bed has been described. The powerful test bed is integrated in the CREW platform. For our research of localization algorithms,

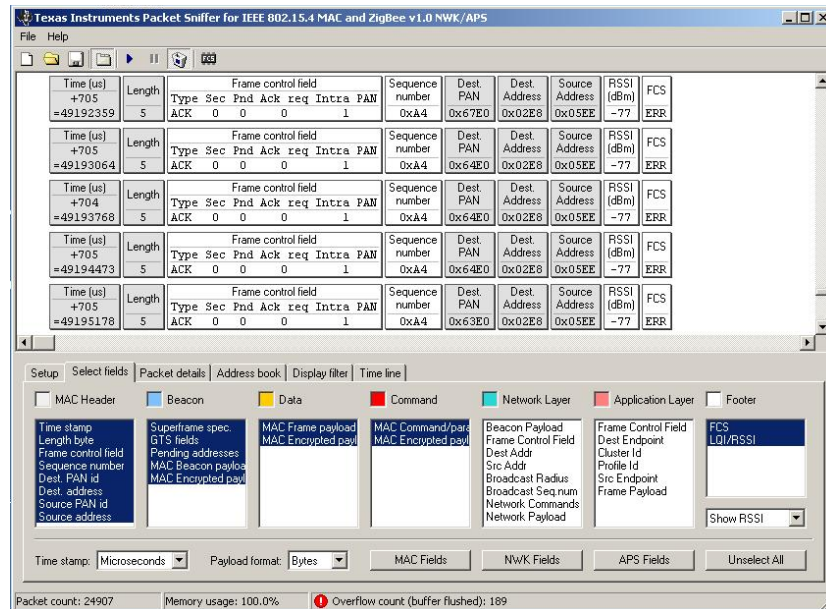


Figure 2.30: A Chipcon packet sniffer detects the generated ACK-signal.

described in chapter 4 and chapter 5, this w-ilab.t test bed has been indispensable.

In a next section, the challenges of working in a wireless environment are dedicated. It is shown that multipath fading is a very annoying phenomenon for indoor localization algorithms. Therefore, a good localization algorithm must cope with the effects of multipath fading. Our experiments show that RSSI-measurements in a long corridor are far too strong (35 dB is not exceptional), due to this (constructive) multipath fading. The section continues with an interference study. Minimum distance between the interferer and an IEEE 802.15.4 DSSS receiver are presented. When two different technologies are housed in the same package, interference can only be avoided using a power reduction of the interferer and a kind of clear channel assessment. The radiation pattern of the antenna is also treated in this section. The radiation pattern of the build in planar inverted-F antenna is quite well omnidirectional, except for two dips. The effect of these dips will be discussed in chapter 4 and chapter 5.

This work aims to improve localization by incorporating physical layer concepts. For fully understanding the pitfalls of the underlying IEEE 802.15.4 DSSS standard, a software defined transmitter has been designed, consisting of a standard PC, LabVIEW, a data acquisition card and a self-made up-converter. Our design meets the IEEE 802.15.4 DSSS physical layer specifications.

References

- [1] *Project Overview: The CREW project*. <http://www.crew-project.eu/>.
- [2] L. Tytgat, B. Jooris, P. De Mil, B. Latré, I. Moerman, and P. Demeester. *Demo Abstract: WiLab, a real-life Wireless Sensor Test Bed with Environment Emulation*. In EWSN: Proceedings of the 6th European Conference on Wireless Sensor Networks, Cork, Ireland, February 11–13 2009.
- [3] S. Bouckaert, W. Vandenberghe, B. Jooris, I. Moerman, and P. Demeester. *The W-iLab.t Testbed*. In Proceedings of the 6th international ICST Conference on Testbeds and Research Infrastructures for the development of Networks & Communities, Berlin, Germany, May 18–20 2010.
- [4] J. Boeckling and T. Van Nieulande. *Ontwerp en Ontwikkeling van Hardware voor een Sensortestbed*. Master's thesis, University College Ghent, 2008.
- [5] P. De Mil, B. Jooris, L. Tytgat, R. Catteuw, I. Moerman, P. Demeester, and A. Kamerman. *Design and Implementation of a Generic Energy-Harvesting Framework Applied to the Evaluation of a Large-Scale Electronic Shelf-Labeling Wireless Sensor Network*. EURASIP Journal on Wireless Communications and Networking, 2010:34369038, 2010.
- [6] *Iperf, netperf, and nuttcp*. <http://staff.psc.edu/benninge/networking/perftests.html>.
- [7] *Iperf - The TCP/UDP Bandwidth Measurement Tool*. <http://iperf.fr/>.
- [8] P. Barsocchi. *Channel Models for Terrestrial Wireless Communications: a Survey*. In CNR-ISTI Technical Report, April 2006.
- [9] T.S. Rappaport. *Wireless Communications: Principles and Practice, Second Edition*. Prentice Hall, 2002.
- [10] J.G. Proakis. *Digital Communications, Fourth Edition*. MacGraw Hill, New York, USA, 2001.
- [11] E. L. Lehmann and G. Casella. *Theory of Point Estimation*. Springer, New York, USA, 1998.
- [12] J. Durbin and S. J. Koopman. *Time Series Analysis by State Space Methods*. Oxford University Press, New York, USA, 2012.

- [13] T. Sarkar, Z. Ji Z, K. Kim KJ, A. Medouri, and M. Salazar-Palma. *A Survey of Various Propagation Models for Mobile Communication*. IEEE Antennas and Propagation Magazine, 45:51–82, 2003.
- [14] *IEEE Standard for Information Technology – Telecommunications and Information Exchange between Systems – Local and Metropolitan Area Networks – Specific Requirements – Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs)*. <http://standards.ieee.org/getieee802/download/802.15.4-2006.pdf>, 2006. IEEE802.15.4-2006 Std.
- [15] E. B. Hamida and G. Chelius. *Investigating the Impact of Human Activity on the Performance of Wireless Networks – An Experimental Approach*. In WoWMoM: Proceedings of the 2010 IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks, pages 1–8, Montreal, Canada, June 14–17 2010.
- [16] Texas Instruments. *CC2420 datasheet*. <http://www.ti.com/lit/ds/symmlink/cc2420.pdf>.
- [17] N. Yarkoni and N. Blaunstein. *Prediction of Propagation Characteristics in Indoor Radio Communication Environments*. PIER 59: Progress in Electromagnetics Research, 59:151–174, 2006.
- [18] K. Fujimoto and J. R. James. *Mobile antenna systems handbook*. Artech House, 2001.
- [19] *IEEE Recommended Practice for Information Technology – Telecommunications and Information Exchange between Systems – Local and Metropolitan Area Networks – Specific Requirements – Part 15.2: Coexistence of Wireless Personal Area Networks with Other Wireless Devices Operating in Unlicensed Frequency Band*. <http://standards.ieee.org/getieee802/download/802.15.2-2003.pdf>, 2003. IEEE802.15.2-2003 Std.
- [20] *802.15.4 CCA-Methods*. http://grouper.ieee.org/groups/802/15/pub/2001/Nov01/01522r0P802-15_TG4-802-15-4-CCA-Methods.doc.
- [21] L. Tytgat, O. Yaron, S. Pollin, I. Moerman, and P. Demeester. *Avoiding Collisions between IEEE 802.11 and IEEE 802.15.4 through Coexistence Aware Clear Channel Assessment*. EURASIP Journal on Wireless Communications and Networking, 2012:137, April 2012.

- [22] *IEEE Standard for Information Technology – Telecommunications and Information Exchange between Systems – Local and Metropolitan Area Networks – Specific Requirements – Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, 2007. IEEE802.11-2007 Std.
- [23] *The Overlapping Channel Problem*. <https://community.ja.net/library/advisory-services/overlapping-channel-problem.pdf>.
- [24] M. Burton. *Channel Overlap Calculations for 802.11b Networks*. In Cirond Technologies White Papers, 2002.
- [25] Silicon Laboratories. *2.4GHz ZigBee Development Kit User Guide*. <http://www.mouser.com/catalog/specsheets/0900766b8078f3b5.pdf>.
- [26] Silicon Laboratories. *Ember ZigBee Development Tools*. <http://www.silabs.com/products/wireless/zigbee/Pages/zigbee-development-tools.aspx>.
- [27] *IEEE Standard for Information Technology – Telecommunications and Information Exchange between Systems – Local and Metropolitan Area Networks – Specific Requirements – Part 15.1: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Wireless Personal Area Networks (WPANs)*. <http://standards.ieee.org/getieee802/download/802.15.1-2005.pdf>, 2005. IEEE 802.15.1 Std.
- [28] F. Vanheel, J. Verhaever, and I. Moerman. *Spectral Interference Study of Wi-Fi on Wireless Sensor Networks*. In Ghent University Faculty of Engineering and Architecture PhD Symposium, page 108, Ghent, Belgium, December 5 2007.
- [29] F. Amoroso. *The Bandwidth of Digital Data Signals*. IEEE Communications Magazine, 18:6:13–24, 1980.
- [30] *IEEE Standard for Information Technology – Telecommunications and Information Exchange between Systems – Local and Metropolitan Area Networks – Specific Requirements – Part 15.3: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for High Rate Wireless Personal Area Networks (WPANs)*. <http://standards.ieee.org/getieee802/download/802.15.3-2003.pdf>, 2003. IEEE 802.15.3-2003 Std.
- [31] S. Bouckaert. *Design and Evaluation of an Auto-Configuring Wireless Mesh Network Architecture*. PhD thesis, Ghent University, 2010.

- [32] *RadioPerf tool*. <http://www.crew-project.eu/portal/wilab/radioperf-tool>.
- [33] Qosmotec. *Air Interface Simulator AIS*. <http://www.qosmotec.com/products/ais/ais-details/26-ais-details>.
- [34] M. Petrova, J. Riihijarvi, P. Mahonen, and S. Laellla. *Performance Study of IEEE 802.15.4 using Measurements and Simulations*. In WCNC 2006: Wireless Communications and Networking Conference, pages 487–492, Las Vegas, NV, USA, April 2006.
- [35] *ZigBee and Wireless Radio Frequency Coexistence*. <https://docs.zigbee.org/zigbee-docs/dcn/07-5219.PDF>.
- [36] C. Hoffmann and P. Russer. *Time-Domain Emission Measurements in K-Band*. *Frequenz*, 65:209–215, August 2011.
- [37] Y. Zhao, B. G. Agee, and J. H. Reed. *Simulation and Measurement of Microwave Oven Leakage for 802.11 WLAN Interference Management*. In Proceedings of the IEEE International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications, Beijing, China, August 8–12 2005.
- [38] H. Kako, T. Nakagawa, and R. Narita. *Development of Compact Inverter Power Supply for Microwave Oven*. *IEEE Transactions on Consumer Electronics*, 37:3:611–616, August 1991.
- [39] M. Nassar, X. E. Lin, and B. L. Evans. *Stochastic Modeling of Microwave Oven Interference in WLANs*. In Proceedings of the IEEE International Conference on Communications, Kyoto, Japan, June 5–9 2011.
- [40] S. Srikanteswara and C. Maciocco. *Interference Mitigation Using Spectrum Sensing*. In Proceedings of the International Conference on Computer Communications and Networks, pages 39–44, Honolulu, Hawaii, USA, August 13–16 2007.
- [41] A. Andersen. *Design Note DN0007: 2.4GHz Inverted F Antenna*. <http://www.ti.com/lit/an/swrul20b/swrul20b.pdf>.
- [42] S. Fang, S. Yeh, and K. Wong. *Planar Inverted-F Antennas for GSM/DCS Mobile Phones and Dual ISM-Band Applications*. In Proceedings of the 2002 IEEE Antennas and Propagation Society International Symposium, volume 4, pages 524–527, San Antonio, TX, USA, June 16–21 2002.
- [43] A. R. Razali and M. E. Bialkowski. *Super Slim Multiband Inverted-F Antenna for GSM/DCS/PCS Operation*. In APMC 2010: Proceedings of the

Asia-Pacific Microwave Conference, pages 227–230, Yokohama, Japan, December 7–10 2010.

- [44] N. A. Saidatul, A. A. H. Azremi, R. B. Ahmad, P. J. Soh, and F. Malek. *A development of Fractal PIFA (planar inverted F antenna) with bandwidth enhancement for mobile phone application*. In Antennas and Propagation Conference, pages 113–116, Loughborough, UK, November 16–17 2009.
- [45] Moteiv Corporation. *Tmote Sky Datasheet*. <http://www.eecs.harvard.edu/~konrad/projects/shimmer/references/tmote-sky-datasheet.pdf>.
- [46] P. Rubbrecht. *Positiebepaling met Draadloze Sensornetwerken in een Anechoïsche Kamer*. Master's thesis, University College Ghent, 2011.
- [47] K. L. Wong and W. J. Chen. *Small-size microstrip-coupled printed PIFA for 2.4/5.2/5.8 GHz WLAN operation in the laptop computer*. Microwave and Optical Technology Letters, 51:9:2072–2076, September 2009.
- [48] T. He, C. Huang, B. M. Blum, J. A. Stankovic, and T. Abdelzaher. *Range-Free Localization Schemes for Large Scale Sensor Networks*. In MobiCom 2003, San Diego, CA, USA, September 14–19 2003.
- [49] GNU Radio. <http://gnuradio.org/redmine/projects/gnuradio/wiki>.
- [50] S. Jaques and K. Tanssens. *Ontwerp van Hard- & Software voor ISM Zender*. Master's thesis, University College Ghent, 2009.
- [51] J. Verhaevert, F. Vanheel, and P. Van Torre. *On the Design of Software and Hardware for a WSN Transmitter*. In 16th Annual Symposium of the IEEE/CVT, pages 11:1–6, Louvain-La-Neuve, Belgium, November 19 2009.
- [52] National Instruments. *LABVIEW*. <http://www.ni.com/labview>.
- [53] National Instruments. *PCI-6110/6111*. <http://www.ni.com/pdf/manuals/370980a.pdf>.
- [54] Analog Devices. *AD8349*. http://www.analog.com/static/imported-files/Data_Sheets/AD8349.pdf.
- [55] D. Banerjee. *PLL Performance, Simulation and Design*. Dog Ear Publishing, 2006.
- [56] R. E. Best. *Phased-Locked Loops, Theory, Design and Applications*. McGraw-Hill, 1993.

-
- [57] Crystek Microwave. *CVCO55BE-2400-2670, Voltage Controlled Oscillator*. <http://www.crystek.com/microwave/admin/webapps/welcome/files/vco/CVCO55BE-2400-2670.pdf>.
- [58] Analog Devices. *ADF4113, RF PLL Frequency Synthesizers*. http://www.analog.com/static/imported-files/data_sheets/ADF4113HV.pdf.
- [59] Chipcon. *Quick Start Instructions CC2420 Development Kit*. <http://www.ti.com/lit/ug/swru044/swru044.pdf>.

3

Localization

3.1 Introduction

The previous chapter discussed the underlying physical aspects of RSSI-based wireless sensor networks. It is shown that a RSSI-based ranging technique is an ideal candidate to implement excellent localization systems (see table 1.5). This chapter presents some simple localization algorithms. In a first step the RSSI-measurements are converted to distances. A second step estimates the actual position using these distances. Section 2.3 demonstrated that multipath fading is by far the most troublesome phenomenon for indoor localization. Indeed, this fading can influence the RSSI-measurements so much that the presence of these outliers makes it very difficult to estimate the position correctly. The mitigation of measurement outliers in localization has seldom been addressed. Therefore, this chapter focuses on the resilience to outliers in widely used techniques that convert the individual calculated distances (ranges) to a position. Geometric multilateration, statistical multilateration, maximum likelihood and a min-max localization algorithm are compared here under the same (extreme) multipath fading conditions. This chapter further emphasizes the importance of keeping the relation between a localization algorithm and the underlying physics.

3.2 Range-based localization steps

Basically, there are two steps in a localization algorithm. The first step converts the physical RSSI-variable to a range representing the distance between an anchor (node knowing its own position) and the target. In a next step, the position is estimated from the ranges.

3.2.1 Conversion of RSSI to a distance

For conversion of a RSSI-measurement to a distance, a channel model is needed. This model characterizes the attenuation of the medium (path loss) between the transmitter and the receiver. The IEEE802.15.4 standard [1] recalculates this path loss in the 2.4 GHz industrial, scientific and medical (ISM) radio band to equation 3.1. The free space model suggests a path loss proportional to the square of the distance; this is equivalent to the first case of equation 3.1. For larger distances the attenuation increases more rapidly than the square of the distance.

$$PL(d) = \begin{cases} 40.2 + 20 \log_{10}(d) & \text{for } d < 8\text{m} \\ 58.5 + 33 \log_{10}(\frac{d}{8}) & \text{for } d > 8\text{m} \end{cases} \quad (3.1)$$

The RSSI-measurements of the anchors are converted to distances, starting from equation 3.1. When the mobile target transmits at 0 dBm, this equation directly yields:

$$\text{RSSI(dBm)} = \begin{cases} -40.2 - 20 \log_{10}(d) & \text{for } d < 8\text{m} \\ -58.5 - 33 \log_{10}(\frac{d}{8}) & \text{for } d > 8\text{m} \end{cases} \quad (3.2)$$

Taking the receiver sensitivity of -92 dBm and a transmit power of +0 dBm, this equation leads to a maximum range of 82.8 m. This justifies the appraisalment “large extended range” in table 1.5, because this is sufficient for most indoor environments. It is a common practice to parameterize this relationship to

$$\text{RSSI(dBm)} = -(A + 10n \log_{10}(d)) \quad (3.3)$$

where the initial signal strength A describes the absolute value of the RSSI, measured at 1 m distance to the transmitting unit. The signal propagation coefficient n shows the attenuation of the signal. For the theoretical approach in the IEEE 802.15.4 standard (see equation 3.2), A equals 40.2 and 58.5 for distances smaller or greater than 8 m respectively and n equals 2 and 3.3 respectively. In practical cases, however both parameters must be determined empirically [2]. Some algorithms choose for “simplicity” among the design goals (see section 1.3) and keep both A and n constant for all anchors. In chapter 4 we present an algorithm where not only the “simplicity” but also the “accuracy” objective is met using basic statistical tools: in this case both A and n are made anchor dependent.

3.2.2 Conversion of distances to a position

The conversion of distances (between the anchors and the target) to a position is a very important step: it shouldn't be too complex. Simple localization algorithm candidates are:

- 1 Multilateration
- 2 Maximum likelihood
- 3 Min-max localization

In the next sections, special attention is given to the effect of outliers on these simple localization algorithms. Accuracy and robustness are also equally important for a valuable localization system (see section 1.3).

3.3 Multilateration

Literation is a simple localization algorithm using RSSI-measurements. Multilateration is twofold: geometric multilateration and statistical multilateration [3, 4].

3.3.1 Geometric multilateration

3.3.1.1 The geometric multilateration principle

Geometric multilateration is based on geometric concepts. The System-on-Chip CC2431 is an example of geometric multilateration with a hardware built-in localization engine [5]. This system obtains an average accuracy of over 4 meters in a deployment area of 10 m x 22 m [6].

In its simplest form geometric multilateration uses three anchors and the geometry of circles, triangles and centroids. This reduces the geometric multilateration to a geometric trilateration. Figure 3.1 presents the ideal case where all distances are estimated correctly. Like in chapter 2 the anchors are drawn in small green circles. The remainder of this trilateration section will use small red circles to show the estimated position. The solid blue lines are reserved for the distances, obtained by equation 3.2. In this ideal case the distance circles intersect in one single point (which is the target).

In figure 3.2 the RSSI-measurements of anchors 2 and 3 are correct, just as in the ideal case. The RSSI-measurement of anchor 1, however is too low. This results in a too large distance circle for this anchor. The (a) part of this figure shows a global view. In the (b) part a detailed view is given. The intersection points A, B and C (represented by small squares) of the circles enclose a polygon. This polygon is the common intersection of the three distance disks. The (b) part further illustrates the detailed principle of the algorithm: the estimated position is

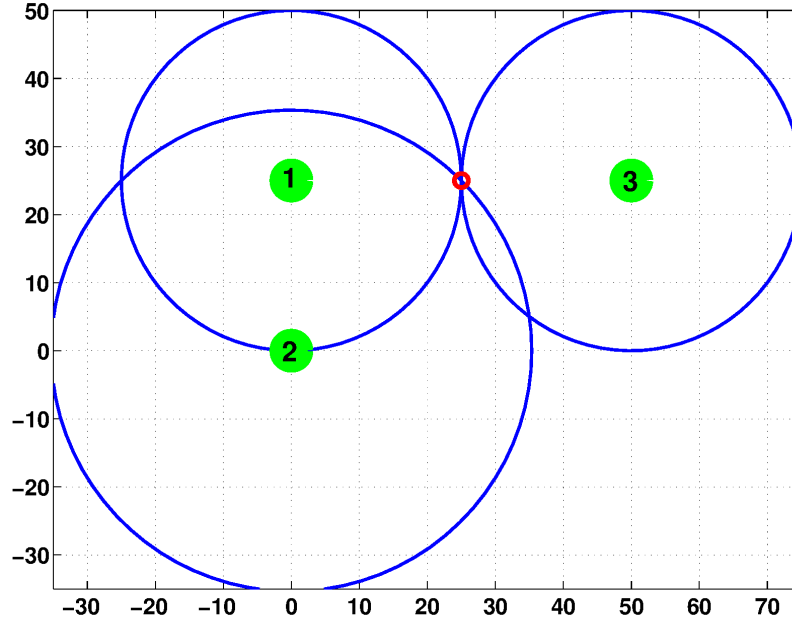


Figure 3.1: Geometric trilateration where all distance circles are correct. The distance circles intersect in one single point: the estimated position, represented by the small red circle.

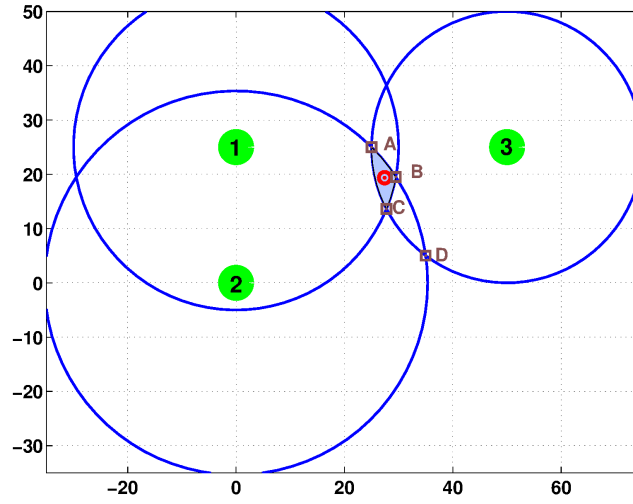
the centroid of the triangle formed by the A, B and C points. The coordinates of this position can easily be calculated by averaging the coordinates of the A, B and C points. For clarity the geometric medians (represented by red dashed lines) of the triangle are also shown.

Geometric multilateration breaks down in many trilaterations. The algorithm starts with a three-anchor-combination of the total anchor set and proceeds with the calculation of the respective centroids. Therefore, in a multilateration algorithm the computational cost rises with the number of anchors [7].

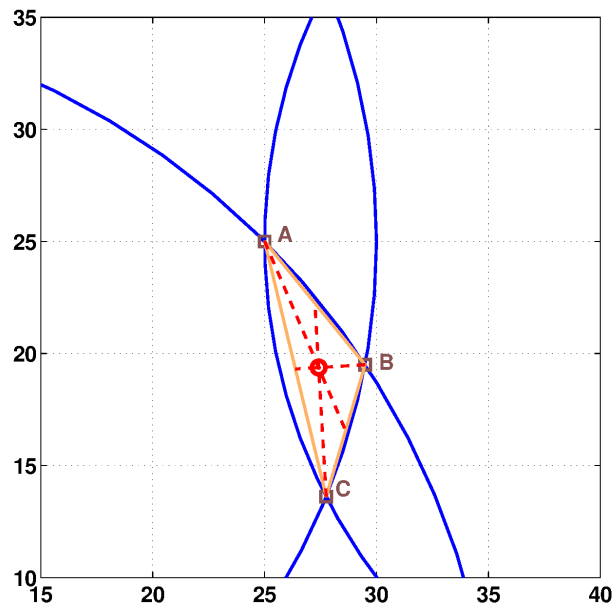
Figure 3.3 shows the number of combinations versus the number of anchors on a logarithmic plot. The number of three-anchor-combinations are represented by the small circles and follow equation 3.4.

$$\#combinations = \binom{\#anchors}{3} = \frac{\#anchors!}{(\#anchors - 3)!3!} \quad (3.4)$$

E.g. in a multilateration algorithm with 12 anchors, there are $\binom{12}{3} = 220$ possible combinations. It can easily be verified that equation 3.4 can be simplified to



(a)



(b)

Figure 3.2: Illustration of the geometric trilateration principle. In (a) a global view is given, the detailed view is shown in (b). The estimated position is the centroid of the triangle with cornerpoints A, B and C. These points are the intersections of the distance circles.

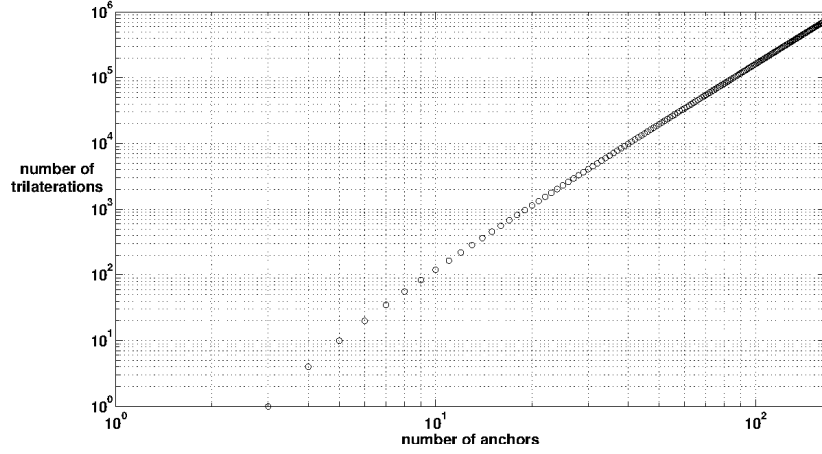


Figure 3.3: With geometric multilateration the required number of trilaterations increases cubically with the number of anchors.

equation 3.5.

$$\#combinations = \frac{\#anchors^3 - 3 \times \#anchors^2 + 2 \times \#anchors}{6} \quad (3.5)$$

Hence, the required number of trilaterations increases cubically with the number of anchors. Next, the results of several trilateration realizations are combined by either averaging the multiple position estimates or by finding their centroid [3]. For each combination, a trilateration is performed and the positions are averaged for the final position estimate.

3.3.1.2 Geometric trilateration with obstructed path outliers

The measured RSSI is further decreased and an outlier is created. The distance circle contains point D in figure 3.2 (a). This point is the other intersection of the anchor 2 and anchor 3 distance circles.

Figure 3.4 reveals that the common intersection of the three distance disks has only two vertices. Indeed, the distance circle of the first anchor is so large that the B-point and the C-point are no longer part of the ‘common’ intersection of the distance disks. This situation is typical for too large distance circles. The triangle of figure 3.2 degrades to the single line AD. This makes the implementation of this algorithm more complex: the centroid of a line differs from the centroid of a triangle.

Possible causes of this outlier include

- destructive multipath fading (see section 2.3.2)

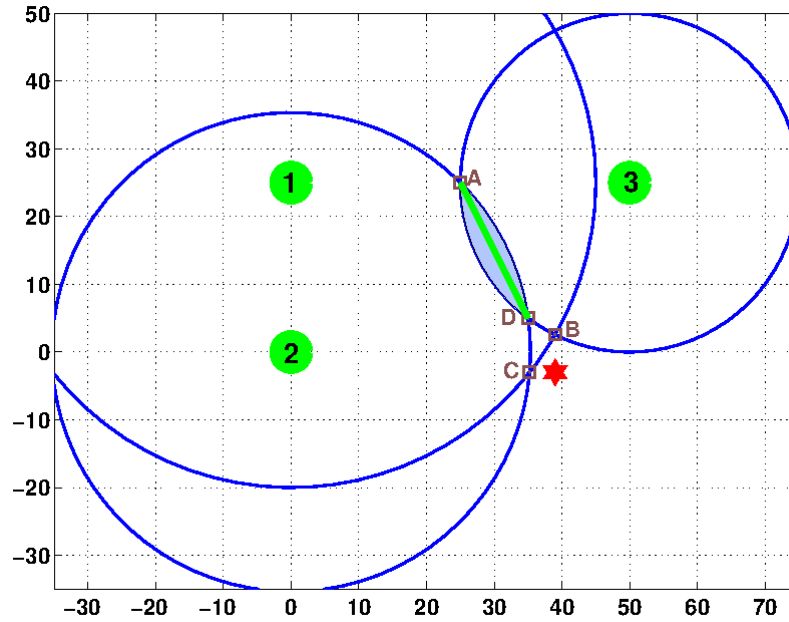


Figure 3.4: The geometric trilateration with an obstructed path outlier. The triangle degenerates to the single line AD. The result of the linearized implementation of the trilateration (red hexagram), see section 3.3.2.2, is not correct: it falls outside the common intersection polygon with vertices A and D.

- the target and the particular anchor connect in a direction with an antenna dip (see section 2.3.4)
- the direct path is obstructed

3.3.1.3 Geometric trilateration with constructive multipath outliers

When a path between an anchor and the target suffers from constructive multipath fading, the measured RSSI is too large (see section 2.3.2). This results in too small distance circles (see equation 3.2). Figure 3.5 illustrates a constructive multipath fading between the anchor 1 and the target. Because no common intersection of the three distance discs can be found, the algorithm can not be applied successfully.

3.3.1.4 A linearized implementation of geometric trilateration

The use of geometric concepts with obstructed path outliers is not straightforward. Under constructive path outlier conditions, no sensible graphical interpretation

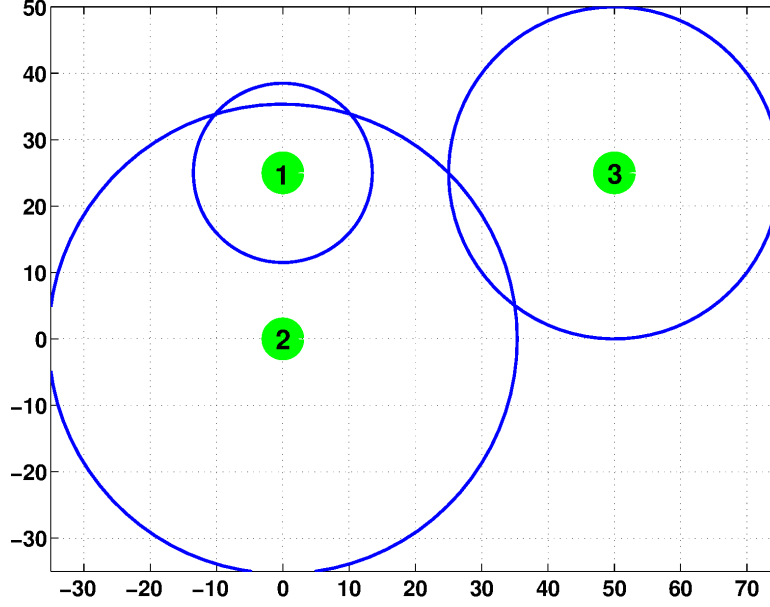


Figure 3.5: The geometric trilateration with a constructive multipath outlier. There is no common intersection of all the distance disks.

can be found. Therefore, the remainder of this section describes a mathematical method that is commonly used for geometric trilateration [8].

Let x and y be the abscissa and ordinate of the (unknown) target. The coordinates of the anchors are represented by (x_i, y_i) with $i=1,2,3$. The distance between the anchor i and the target is denoted by d_i . Hence, 3.6 is a set of nonlinear equations.

$$d_i = \sqrt{(x_i - x)^2 + (y_i - y)^2} \quad (i = 1, 2, 3) \quad (3.6)$$

Squaring both sides of these equations and subtracting the last equation from the other two linearizes the set. In matrix form this can be written as

$$A \times \begin{bmatrix} x \\ y \end{bmatrix} = B \quad (3.7)$$

with $A = 2 \begin{bmatrix} x_3 - x_1 & y_3 - y_1 \\ x_3 - x_2 & y_3 - y_2 \end{bmatrix}$ a 2x2 and $B = \begin{bmatrix} d_1^2 - d_3^2 - x_1^2 - y_1^2 + x_3^2 + y_3^2 \\ d_2^2 - d_3^2 - x_2^2 - y_2^2 + x_3^2 + y_3^2 \end{bmatrix}$ a two-column vector, both with known elements. Hence equation 3.7 represents a linear system with a set of two equations in the two variables x and y . In this trilateration approach both sides of equation 3.7 are left-sided multiplied by the

inverse of matrix A , yielding

$$\begin{bmatrix} x \\ y \end{bmatrix} = A^{-1} \times B \quad (3.8)$$

This confirms what other researchers [3] have found: the geometric trilateration algorithm is simple to implement and computationally inexpensive. The implementation of geometric trilateration via equation 3.8, however, is erroneous for outliers in an obstructed path environment. The red hexagram in figure 3.4, representing the outcome of this equation, is outside the intersection of the distance discs. Furthermore, it shows that the too large distance of anchor 1 is corrected in the wrong sense: the faulty distance further increases.

3.3.2 Statistical multilateration

3.3.2.1 The statistical multilateration principle

The previous section revealed that geometric multilateration is not scalable in the number of anchors. The cubic increase with the number of anchors in equation 3.5 quickly results in an unmanageable amount of trilaterations. Using only part of the abundant measurements might look interesting, but poses other problems: e.g. how to discriminate between a good measurement and an outlier, or how to determine the optimal combination of multiple measurements? In order to achieve the design goals of section 1.3, other localization algorithms are needed. It is widely accepted that statistics can manage large data sets.

In a statistical approach, the variables are not deterministic but are treated as random variables [9] with probability density functions [10]. Equation 3.9 represents a nonlinear set of m equations in the statistical variables x and y .

$$\hat{d}_i = \sqrt{(x_i - x)^2 + (y_i - y)^2} + \eta_i \quad (i = 1, 2, 3 \dots m) \quad (3.9)$$

The estimated distances are denoted by \hat{d}_i , m is the total number of anchors and η_i denotes an additive noise, which is usually assumed to be an independent zero-mean Gaussian distributed random variable. The overdetermined dataset of equation 3.9 can be rewritten as equation 3.10

$$\hat{d}_i = d_i(x, y) + \eta_i \quad (i = 1, 2, 3 \dots m) \quad (3.10)$$

where $d_i(x, y)$ represents the Euclidean distance between anchor i and the position of the target. \hat{d}_i the estimated distance (calculated with the RSSI). Under the Gaussian distribution condition of the previous paragraph, the interpretation of equation 3.10 is straightforward: the estimated distances are normally distributed around the Euclidean distances.

This equation can easily put in matrix form:

$$\hat{\mathbf{d}} = \mathbf{d}(x, y) + \boldsymbol{\eta} \quad (3.11)$$

where $\hat{\mathbf{d}}$ and $\boldsymbol{\eta}$ are multivariate column vectors and $\mathbf{d}(x, y)$ is a nonlinear vector function.

Solving this equation can be done by

- In constrained minimization [11], weighted constraints are added to the least squares problem. The goal is to find the solution, minimizing the weight of the constraints using the Lagrange multiplier [12].
- Multidimensional scaling (MDS) [13] is a statistical technique for analyzing similarities or dissimilarities in data in various fields, ranging from information science over molecular biology, marketing, geography, ecology and sociology to psychology. In [14] it is used in a localization algorithm. Here, the first step is the construction of a scalar distance matrix. Next, a least square algorithm is used to solve a set of linear equations using eigenvalues.
- Subspace decomposition [15] is another technique. It is based on multidimensional scaling.
- The appearance of the error term $\boldsymbol{\eta}$ makes equation 3.9 a “mathematical programming” or “numerical optimization [16]” problem, and more specific a “nonlinear programming [17]” problem. In an initial phase the problem is linearized, using Taylor expansion, gradients, Jacobians and/or Hessians. Further processing can be performed with following iterative techniques:
 - the Steepest descent algorithm
A method like this converges, but the final convergence is linear and often very slow [18]
 - the Conjugate gradient algorithm [19]
 - the Gauss-Newton algorithm
In many applications, the Gauss-Newton has quite good performance, though it normally has only linear convergence [18].
 - the Levenberg-Marquardt algorithm
This method is based on Gauss-Newton. It has superior global performance.
- The algorithms of the previous item require iterations, affecting processing time. It is a common practice [20] to linearize equation 3.9 using the subtraction method. In section 3.3.2.2, this linear least squares algorithm is applied on a trilateration problem.

3.3.2.2 A linearized implementation of statistical trilateration

To illustrate this simple implementation with an example, we start with the extension of the linearization procedure described in section 3.3.1. The set of m equations (3.9) is squared and the last equation is subtracted from each of the others yielding equation 3.12. A is a $(m - 1) \times 2$ and B a $(m - 1)$ column vector, both with known elements.

$$A \times \begin{bmatrix} x \\ y \end{bmatrix} = B \quad (3.12)$$

Because matrix A is non-square, its inverse does not exist. However, the over determined set of linear equations can be solved by the use of a generalized inverse, also called the pseudo-inverse matrix A^+ . A well-known pseudo-inverse matrix is the Moore-Penrose pseudo-inverse [21, 22]. This matrix is the solution of the least square problem: it minimizes the sum of the squared differences between the data values and their corresponding modeled values. When A has full column rank, so that $A^T A$ is invertible, it can be proven that $A^+ = (A^T A)^{-1} A^T$ where A^T denotes the transposed matrix of matrix A . Substituting A^+ in equation 3.12 results in equation 3.13, the solution of the statistical multilateration problem.

$$\begin{bmatrix} x \\ y \end{bmatrix} = (A^T A)^{-1} A^T B \quad (3.13)$$

This brings us to the following discussion:

- Statistical multilateration with the help of equation 3.13 is perfectly scalable in the number of anchors.
- It can easily be verified that $A^+ = A^{-1}$ for a nonsingular, symmetrical matrix. The geometric trilateration in the linear system in equation 3.8 and the statistical trilateration in the linear system in equation 3.13 are therefore identical. The implementation of statistical multilateration with the help of equation 3.13, however does not minimize the sum of the squared differences between the calculated distances and their corresponding modeled values: it minimizes the squared differences between the position and the “transformed” position with the help of the A and B matrices. The transformation introduces squares of the \hat{d}_i in the B -matrix, which makes neither the physical nor the graphical interpretation straightforward. The localization algorithms, described in chapter 4 and chapter 5 avoid this situation.
- This implementation of statistical trilateration is not better than the implementation of geometric trilateration that is discussed in section 3.3.1. Therefore it offers no solution to outliers. The red hexagram in figure 3.4 represents the result of the trilateration. This red hexagram is outside the common

intersection (represented by the polygon with vertices A and D) of the distance disks. Therefore, this linearized implementation is erroneous in the presence of an obstructed path outlier.

- Many authors testify that the quality of the statistical multilateration model increases when there are more anchors. This is in line with our findings because treating more data simultaneously also improves the quality of a regression model, under the explicit condition that these are no outliers. In chapter 4 we will pay extra attention to outliers and base our 2D localization algorithm on the quality of the regression model.
- Statistical methods need a large data set. If this dataset is reduced, the quality of the underlying model deteriorates.

3.4 Maximum likelihood

3.4.1 The maximum likelihood principle

The maximum likelihood algorithm is also a statistical method, and therefore based on equation 3.10. Plotting the calculated distances on a vertical axis versus the modeled distances reveals the graphical interpretation of the errors η_i . It is a common practice to reform equation 3.10 to

$$\sum \eta_i^2 = \sum (\hat{d}_i - d_i)^2 \quad (3.14)$$

The interpretation is straightforward: if all calculated distances are equal to the modeled distances, the model is extremely good. This explains the essence of the minimum mean square error cost function (MMSE):

$$(\hat{x}, \hat{y}) = \underset{(x,y)}{\operatorname{argmin}} \sum_{j \in \operatorname{anchor}(i)} (\hat{d}_{i,j} - d_{i,j})^2 \quad (3.15)$$

(\hat{x}, \hat{y}) represents here the estimated (or most likely position). This position is found at the location where the sum of squared error between the measured position and the modeled position is minimal. This statistical method combines different RSSI-measurements in a joint density function [23]. As an example the maximum likelihood approach is used for the layout of figure 3.1. Figure 3.6 shows the contribution of each anchor in equation 3.15. The green anchor 1 has a large contribution at the left-hand side extreme of the building, the blue and the red at the right-hand side extreme. The summation of the individual contributions results in figure 3.7. For this example, the minimum of this cost function results in a correct position estimate: the target.

Several cost functions exist, depending on the presumed distribution of η_i and the implementation of the model. Indeed, equation 3.11 can be extended to represent other main variables which indirectly depend on the distance e.g. the RSSI.

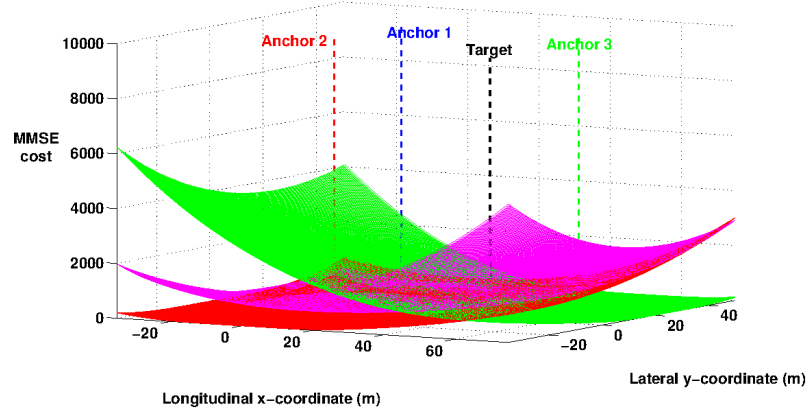


Figure 3.6: The contribution of each anchor in the (joint density) MMSE cost function

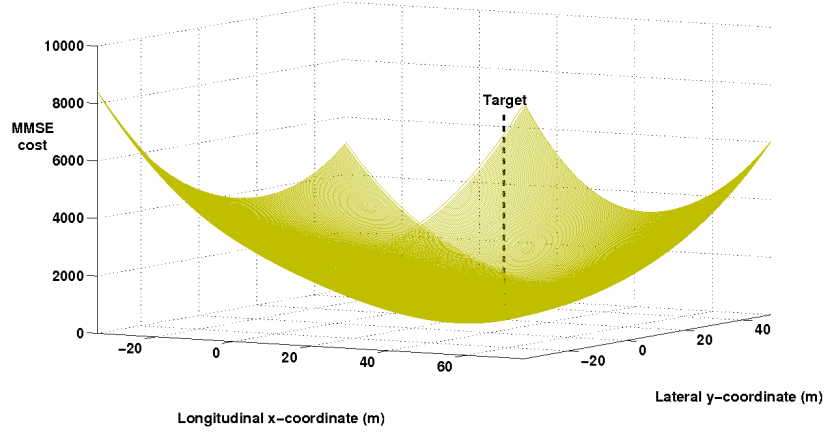


Figure 3.7: The MMSE cost function for the case where all distances are estimated correctly

Under the assumption that the η_i in equation 3.10 are Gaussian and independent identically distributed, the probability distribution of the residuals can be written as:

$$p(\boldsymbol{\eta}) = \prod_{i=1}^m \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\eta_i^2}{2\sigma^2}\right) = (2\pi\sigma^2)^{-m/2} \exp\left(-\frac{\sum_{i=1}^m \eta_i^2}{2\sigma^2}\right) \quad (3.16)$$

This equation 3.16 reveals that maximizing the Gaussian likelihood with respect to the signal parameters is equivalent to minimizing the squares of the residuals [24]. Furthermore, maximizing $p(\boldsymbol{\eta})$ is equivalent to minimizing $-p(\boldsymbol{\eta})$. Therefore, it always needs to be specified whether the cost function needs to be minimized or

maximized. Cost functions are presented in [25–28]. In section 4.5 we will present a new cost function based on the physical aspects (the propagation channel) that needs to be maximized. Because most cost functions are nonlinear, they need to be solved either by the methods described in 3.3.2.1 or by the grid method [29].

3.4.2 False extrema and false convergences

For most cases of practical interest the performance of maximum likelihood algorithms is optimal for sufficiently large data records [29]. This can be explained by the fact that the distribution of η_i approaches a normal distribution, thanks to the central limit theorem. For small datasets and in presence of outliers, this is no longer true. These outliers give rise to local extrema in the cost function, which put high stress on the methods mentioned in section 3.3.2.1: it is possible these methods do not converge any longer or converge to the wrong extreme.

An exacerbated situation is shown in figure 3.8. Minimizing the cost function does not result in a good target localization, the local minimum is by far a better choice. Good models and efficient preprocessing improve this situation [30].

3.5 Min-Max localization algorithm

3.5.1 The Min-Max localization principle

This section illustrates the principle of the Min-Max localization algorithm, starting again (just like in chapter 3.3) with an example where the range measurements are exact. This algorithm [31–33] uses squared bounding boxes, obtained by adding and subtracting the estimated distance from the anchor's position. Let \hat{d}_i be the estimated distance between the target and the anchor position with coordinates (x_i, y_i) . Equation 3.17 describes these bounding boxes, drawn in dashed lines in figure 3.9.

$$[x_i - \hat{d}_i, y_i - \hat{d}_i] \quad \times \quad [x_i + \hat{d}_i, y_i + \hat{d}_i] \quad (3.17)$$

The intersection of these bounding boxes, which is presented by the pink rectangle, is obtained by computing the maximum of all coordinate minima and the minimum of all maxima. Table 3.1 contains the needed information in matrix-form for our example of figure 3.9. In the first column we find the transmitting anchor, the second represents x-coordinate of that anchor, the third gives the y-coordinate, the fourth the estimated distance (with the help of equation 3.2), the fifth and the sixth the coordinates minus the estimated distance and the seventh and eighth the coordinates plus the estimated distance.

Next, the maxima of the fifth and sixth columns are calculated. These define the maxima of the minimum x- and y- coordinate of the intersection respectively.

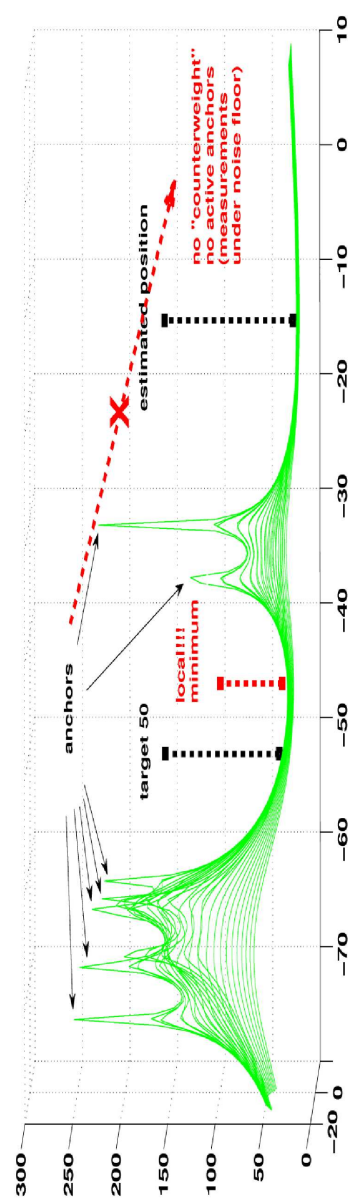


Figure 3.8: A cost function in a real-life environment with both constructive multipath and obstructed path outliers: the local minimum is closer to the target than the absolute minimum [30]

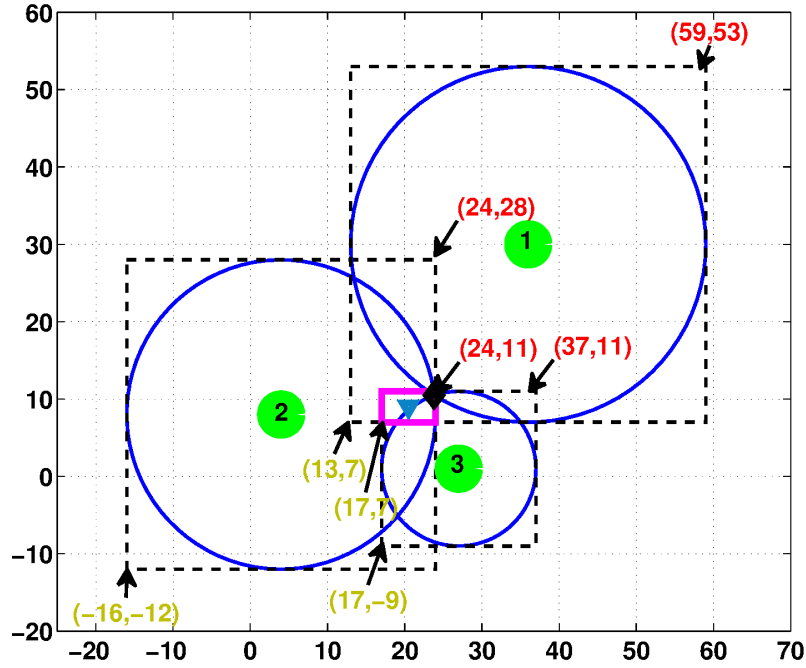


Figure 3.9: Principle of the Min-Max localization algorithm.

i	x_i	y_i	\hat{d}_i	$x_i - \hat{d}_i$	$y_i - \hat{d}_i$	$x_i + \hat{d}_i$	$y_i + \hat{d}_i$
1	36	30	23	13	7	59	53
2	4	8	20	-16	-12	24	28
3	27	1	10	17	-9	37	11
				max=17	max=7	min=24	min=11

Table 3.1: Min-Max algorithm matrix (for the case that all distances are estimated correctly)

The calculation of the minima of the seventh and eighth column follows, resulting in minima of maximum coordinates. The two pairs of coordinates define the intersection of the bounding boxes, mathematically expressed as equation 3.18.

$$[\max(x_i - \hat{d}_i), \max(y_i - \hat{d}_i)] \times [\min(x_i + \hat{d}_i), \min(y_i + \hat{d}_i)] = \quad (3.18)$$

$$[17, 7] \times [24, 11]$$

The intersection can also be found graphically (although this is more complex). The starting point is the (green) marking of the left bottom coordinates of the bounding boxes. The example of figure 3.9 illustrates the navigation from the leftmost (green) lower corner (-16, -12) to the green bounding box corners that

are more to the right-hand side: the intermediate corner is (13,7) and (17,-9) is the green bounding box corner that is the most to the right-hand side. Shifting upwards from this point in the direction of the other green corner(s) determines the lower left-hand side corner (17,7) of the intersection. An analogous procedure determines the upper right-hand side corner of the intersection: from the (red marked) (59,53) toward (37,11) and (24,28) to terminate in (24,11).

The final estimated position, the small blue triangle, is set by calculating the centroid of the pink rectangle. For a rectangle this centroid is simply the average of the corner coordinates. Figure 3.9 further reveals that the estimate position deviates from the target (presented by the black diamond) when all distances are exact. This effect is called biased estimation and this figure shows that in our particular example the error on the position decreases when the path between anchor 2 and the target is slightly obstructed: the RSSI-reading decreases, the distance circle increases (equation 3.2), the bounding box expands for this anchor (equation 3.17) and the small blue triangle approaches the (fixed) target. Removing the bias in an algorithm is a better option than counting on an unwanted phenomenon to improve the accuracy. We present in chapter 4 new algorithms that are based on circles (not boxes) to improve this bias.

The main advantage of the Min-Max algorithm, however, remains: it is fast and can be easily executed on a 8-bit MCU with matrix manipulations that are converted to simple addition, subtraction and comparison operations [33].

3.5.2 Min-Max positioning algorithm with obstructed path outliers

Both an obstructed path and a destructive multipath fading result in too small RSSI- readings. This implies that the estimated distances are too large. This section demonstrates the consequences for a Min-Max localization algorithm under these circumstances. In figure 3.10, the distance circle of anchor 1 is largely increased. The corresponding matrix is shown in table 3.2.

i	x_i	y_i	\hat{d}_i	$x_i - \hat{d}_i$	$y_i - \hat{d}_i$	$x_i + \hat{d}_i$	$y_i + \hat{d}_i$
1	36	30	40	-4	-10	76	70
2	4	8	20	-16	-12	24	28
3	27	1	10	17	-9	37	11
				max=17	max=-9	min=24	min=11

Table 3.2: Min-Max algorithm matrix with an obstructed path outlier

Next, the principle of the previous section 3.5.1 is applied, resulting in the

intersection of equation 3.19.

$$\begin{aligned} & [\max(x_i - \hat{d}_i), \max(y_i - \hat{d}_i)] \times [\min(x_i + \hat{d}_i), \min(y_i + \hat{d}_i)] = \\ & [17, -9] \times [24, 11] \end{aligned} \quad (3.19)$$

The position is estimated at $(\hat{x}, \hat{y}) = (\frac{24+17}{2}, \frac{-9+11}{2}) = (20.5, 1)$. The intersection of the bounding boxes contains the target. The intersecting area, however is much larger than in the previous section 3.5.1, the y-axis bounds are determined by the same anchor (number 3) in this particular example, resulting in a larger error. Despite this large error, the Min-Max algorithm is used to detect and correct the error of a maximum likelihood algorithm [34] in an obstructed path environment. A Min-Max algorithm that is more complex could use the knowledge that bounds are determined by the same anchor. In chapter 4 we present an algorithm that detects and deletes or corrects too large distance circles.

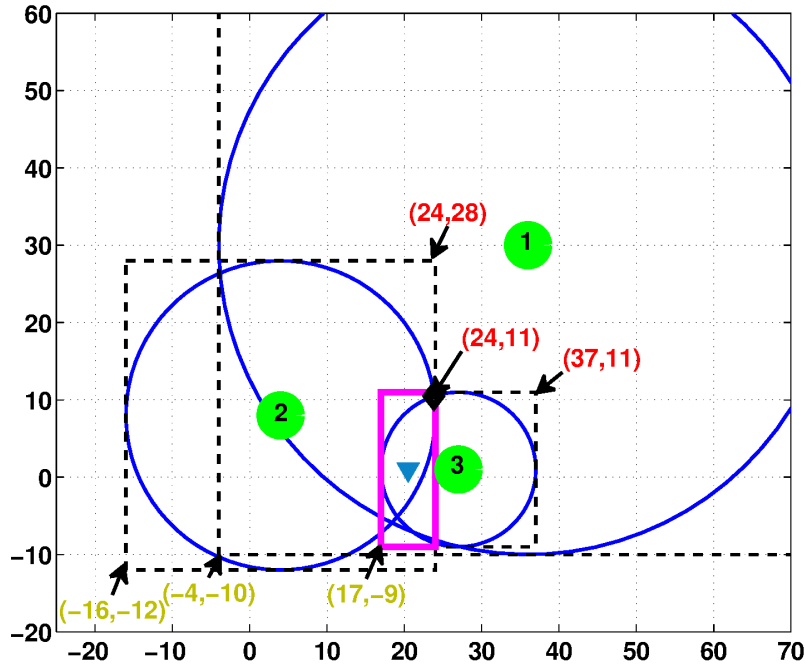


Figure 3.10: The Min-Max localization algorithm with an obstructed path outlier.

3.5.3 Min-Max positioning algorithm with constructive multipath fading

This section illustrates the working of a Min-Max positioning algorithm, when there is an outlier due to constructive multipath fading (see section 2.3.2). In presence of constructive multipath fading, the RSSI-measurements are too high, resulting in too small distance circles. In figure 3.11 the constructive multipath fading between anchor 2 and the target is so high that an outlier is created. Table 3.3 represents the corresponding matrix. Again, the principles of section 3.5.1 are applied, resulting in the bounding box, represented by equation 3.20:

$$\begin{aligned} & [\max(x_i - \hat{d}_i), \max(y_i - \hat{d}_i)] \times [\min(x_i + \hat{d}_i), \min(y_i + \hat{d}_i)] = \\ & [17, 7] \times [10, 11] \end{aligned} \quad (3.20)$$

The position is estimated at $(\hat{x}, \hat{y}) = (\frac{17+10}{2}, \frac{7+11}{2}) = (13.5, 9)$.

A closer look at table 3.3 reveals that the maximum of all x-coordinate minima is larger than the minimum of all x-coordinate maxima. Performing the graphical construction of the “intersection” box again (see section 3.5.1), shows that the green bounding-box lower cornerpoints define a lower right-hand side (not lower left-hand side) box corner. Likewise, the red bounding-box upper corners define an upper left-hand side (not upper right-hand side) box corner of the intersection. To our knowledge, this problem is not described in literature, yet. One solution might be allowing the “negative sides”, like in the figure 3.11. The error, however is large and the intersection box no longer contains the target. Another solution might be ignoring the faulty anchor. Even if it can be identified, the error stays high. We will come back on this discussion, when we treat outliers in our preprocessing and positioning algorithm in chapter 4.

3.6 Case study: Effect of outliers on simple localization algorithms with three anchors

A simulation is set up for comparing the different localization algorithms. We limit this simulation to the usage of no more than three anchors. The first three

i	x_i	y_i	\hat{d}_i	$x_i - \hat{d}_i$	$y_i - \hat{d}_i$	$x_i + \hat{d}_i$	$y_i + \hat{d}_i$
1	36	30	23	13	7	59	53
2	4	8	6	-2	2	10	14
3	27	1	10	17	-9	37	11
				max=17	max=7	min=10	min=11

Table 3.3: Min-Max algorithm matrix with a constructive multipath outlier

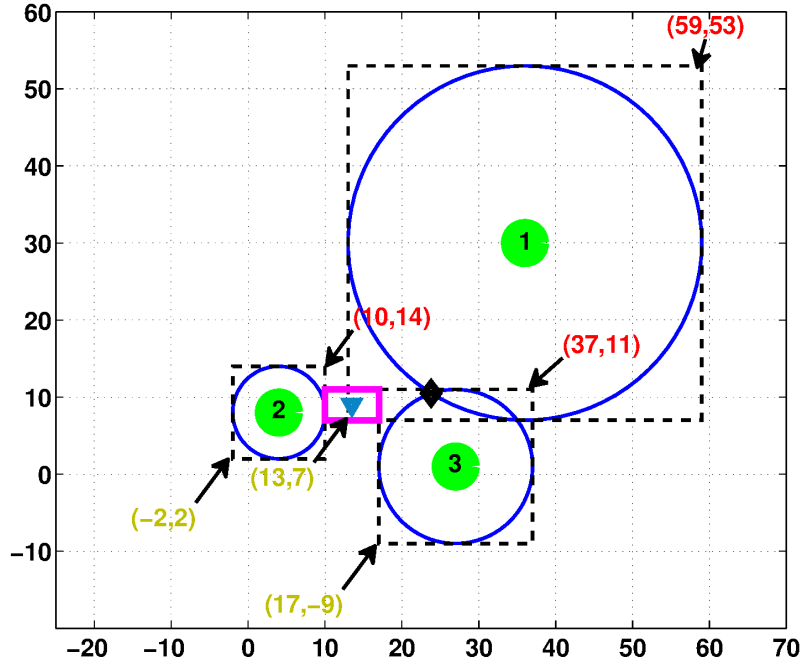


Figure 3.11: The Min-Max localization algorithm with a constructive multipath outlier.

columns of table 3.4 identify the anchors and their fixed position. The fourth column shows the respective estimated distances: the second and third anchor distances are estimated correctly. The estimated distance for the first anchor is changed from 13.75 m to 45 m. The last column refers to the respective situation (changing from constructive multipath outlier to obstructed path outlier) and the corresponding subplot in figure 3.12.

i	x_i	y_i	\hat{d}_i	illustration
1	0	25	13.75	constructive multipath outlier, see figure 3.12 (e)
			20	constructive multipath, see figure 3.12 (c)
			25	exact, see figure 3.12 (a)
			30	obstructed path, see figure 3.12 (b)
			45	obstructed path outlier, see figure 3.12 (d)
2	0	0	70.7	
3	50	25	25	

Table 3.4: Outlier test setup for comparison of simple localization algorithms

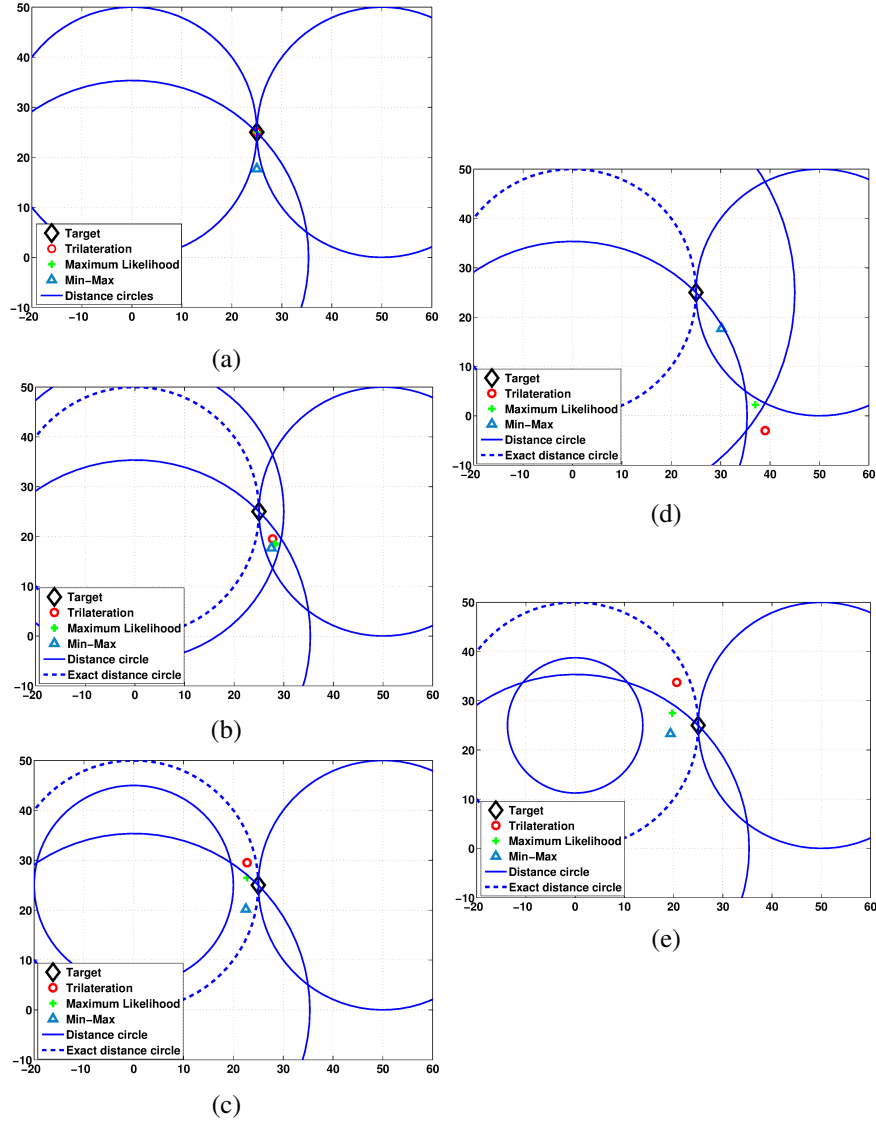


Figure 3.12: Comparison of different localization algorithms with three anchors for (a) correct estimated distances, (b) an obstructed path environment, (c) a constructive multipath environment, (d) an outlier obstructed path environment, and (e) an outlier constructive multipath environment.

In this figure the target is presented by a black diamond, the solid blue lines denote the calculated distance and the dashed blue lines are for the exact distance between the target and anchor 1 (25 m).

The small red circles show the position estimation of the linearized trilateration algorithm, the green “+”-signs are for the MMSE maximum likelihood algorithm and the blue triangles are for the Min-Max algorithm.

Figure 3.12(a) shows that only the Min-Max algorithm is biased. The errors of this algorithm are more constant for the different cases (a) through (e). The linearized trilateration and MMSE maximum likelihood show the same pattern: no error for the exact distance and increasing for the other cases in the same order. The worst accuracy is met for the disturbance with an obstructed path outlier for both algorithms. This confirms the correctness of using a Min-Max algorithm to correct errors of a maximum likelihood algorithm in an obstructed path environment, like in [34].







	Accuracy	Robustness to outliers
Linearized Trilateration		
Maximum likelihood		
Min-Max		

Table 3.5: Simple localization algorithms with three anchors versus design goals.

Table 3.5 gives an overview of the results of this case study. The accuracy score is based on the results under minimal distance errors, see figure 3.12 (a), (b) and (c). The robustness to outliers score is based on the results under severe distance disturbances, see figure 3.12 (d) and (e).

The highest accuracy is for the maximum likelihood algorithm. Second is the linearized trilateration algorithm. The Min-Max algorithm is the least susceptible to outliers.

3.7 Conclusions

This chapter started with the presentation of some common simple localization algorithms. A comparison between a linearized trilateration algorithm, a Min-Max algorithm, and a Maximum likelihood algorithm has been made. There is no doubt that there is a trade-off between complexity, accuracy, and robustness to outliers.

Simple localization algorithms, however can perform excellently if some pitfalls are avoided:

- A good localization algorithm keeps the underlying physical aspects transparent in order to avoid faulty implementations.
- A good preprocessing technique is needed to eliminate outliers.
- Furthermore, statistical methods enable an accurate manipulation of a large data set.

In the next chapter, we present localization algorithms that fulfill these requirements.

References

- [1] *IEEE Standard for Information Technology – Telecommunications and Information Exchange between Systems – Local and Metropolitan Area Networks – Specific Requirements – Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs)*. <http://standards.ieee.org/getieee802/download/802.15.4-2006.pdf>, 2006. IEEE802.15.4-2006 Std.
- [2] A. Wessels, X. Wang, R. Laur, and W. Lang. *Dynamic Indoor Localization using Multilateration with RSSI in Wireless Sensor Networks for Transport Logistics*. In Proceedings of the XXIV Conference on EuroSensors, pages 220–223, Linz, Austria, September 5–8 2010.
- [3] D. Munoz, F. Bouchereau, C. Vargas, and R. Enriquez. *Position Location Techniques and Applications*. Academic Press, Burlington, MA, USA, 2009.
- [4] P. Rubbrecht. *Positiebepaling met Draadloze Sensornetwerken in een Anechoïsche Kamer*. Master’s thesis, University College Ghent, 2011.
- [5] Texas Instruments. *System_on_Chip for 2.4 GHz ZigBee(TM)/IEEE802.15.4 with Location Engine*. <http://www.ti.com/product/cc2431>, 2007.
- [6] G. V. Merrett, A.S. Weddell, L. Berti, N. R. Harris, N. M. White, and B. M. Al-Hashimi. *A Wireless Sensor Network for Cleanroom Monitoring*. In EuroSensors 2008, pages 1553–1556, Dresden, Germany, September 7–11 2008.
- [7] C. Shih and P.J. Marron. *COLA: Complexity-reduced Trilateration Approach for 3D Localization in Wireless Sensor Networks*. In SENSORCOM: Fourth International Conference on Sensor Technologies and Applications, pages 24–32, Venice, Italy, July 18–25 2010.
- [8] F. Thomas and L. Ros. *Revisiting Trilateration for Robot Localization*. IEEE Transactions on Robotics, 21:1:93–101, February 2005.
- [9] J. S. Bendat and A. G. Piersol. *Random Data, Analysis and Measurement Procedures*. John Wiley & Sons Academic Press, New York, USA, 1986.
- [10] D. J. Torrieri. *Statistical Theory of Passive Location Systems*. IEEE Transactions on Aerospace and Electronic Systems, 20:2:183–198, March 1984.

- [11] K. W. Cheung, H. C. So, W. K. Ma, and Y. T. Chan. *Least Squares Algorithms for Time-of-arrival-based Mobile Location*. IEEE Transactions on Signal Processing, 52:4:1121–1130, April 2004.
- [12] D. P. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods*. Athena Scientific, Nashua, NH, USA, 1996.
- [13] J. B. Kruskal and M. Wish. *Multidimensional Scaling*. Sage Publications, Newbury Park, CA, USA, 1978.
- [14] H.-W. Wei, Q. Wan, Z.-X. Chen, and S.-F. Ye. *A Novel Weighted Multidimensional Scaling Analysis for Time-of-arrival-based Mobile Location*. IEEE Transactions on Signal Processing, 56:7:3018–3022, July 2008.
- [15] H. So and F. Chan. *A Generalized Subspace Approach for Mobile Positioning with Time-of-arrival Measurements*. IEEE Transactions on Signal Processing, 55:10:5103–5107, October 2007.
- [16] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, New York, USA, 1999.
- [17] M. S. Bazaraa, H. D. Sherali, and C. M. Shetti. *Nonlinear Programming: Theory and Algorithms*. Wiley-Interscience, Hoboken, NJ, USA, 2006.
- [18] K. Madsen, H. B. Nielsen, and O. Tingleff. *Methods for Non-Linear Least Squares Problems*. Technical Report IMM2004-03215, Informatics and Mathematical Modelling, Technical University of Denmark, April 2004.
- [19] J. L. Nazareth. *Conjugate-gradient Methods*. Wiley Interdisciplinary Reviews: Computational Statistics, 1:348–353, 2009.
- [20] I. Güvenc, S. Gezici, F. Watanabe, and H. Inamura. *Enhancements to Linear Least Squares Localization Through Reference Selection and ML Estimation*. In WCNC 2008: Wireless Communications and Networking Conference, pages 284–289, Las Vegas, NV, USA, April 2008.
- [21] R. Penrose. *A Generalized Inverse for Matrices*. Proceedings of the Cambridge Philosophical Society, 51:404–413, 1955.
- [22] E. H. Moore. *On the Reciprocal of the General Algebraic Matrix*. Bulletin of the American Mathematical Society, 26:9:394–395, 1920.
- [23] L. L. Scharf. *Statistical Signal Processing Detection, Estimation, and Time Series Analysis*. Addison-Wesley, 1991.
- [24] J. J. K. O Ruanaidh and W. J. Fitzgerald. *Numerical Bayesian Methods Applied to Signal Processing*. Springer, New York, USA, 1996.

- [25] N. Patwari, A. O. Hero, M. Perkins, N. S. Correal, and R. J. O'Dea. *Relative Location Estimation in Wireless Sensor Networks*. IEEE Transactions on Signal Processing, 51:2137–2248, 2003.
- [26] M. Takashima, D. Zhao, K. Yanagihara, K. Fukui, S. Fukunaga, S. Hara, and K. Kitayama. *Location Estimation using Received Signal Power and Maximum Likelihood Estimation in Wireless Sensor Networks*. Electronics and Communications in Japan Part I-Communications, 90:62–72, 2007.
- [27] I. Yamada, T. Ohtsuki, T. Hisanaga, and L. Zheng. *An Indoor Position Estimation Method by Maximum Likelihood Algorithm using RSS*. In Proceedings of SICE Annual Conference, pages 2918–2921, Takamatsu City, Kagama, Japan, September 17-20 2007.
- [28] H. Cho, M. Kang, J. Park, B. Park, and H. Kim. *Performance Analysis of Location Estimation Algorithm in ZigBee*. In Proceedings of the 21st International Conference on Advanced Networking and Applications Workshops/Symposia, pages 302–306, Niagara Falls, Ontario, Canada, May 21-23 2007.
- [29] S. M. Kay. *Fundamentals of Statistical Signal Processing, Volume 1: Estimation theory*. Prentice Hall, Upper Saddle River, NJ, USA, 1993.
- [30] F. Vanheel, J. Verhaevert, E. Laermans, and I. Moerman. *Comparison of 2D RSSI based WSN Multipath Faded Indoor Localization Algorithms*. In Ghent University Faculty of Engineering and Architecture PhD Symposium, page 212, Ghent, Belgium, December 1 2010.
- [31] Q. Q. Shi, H. Huo, T. Fang, and D. R. Li. *A Distributed Node Localization Scheme for Wireless Sensor Networks*. Wireless Personal Communications, 53:1:15–33, March 2010.
- [32] H. Y. Luo, H. Li, F. Zhao, and J. H. Peng. *An Iterative Clustering-Based Localization Algorithm for Wireless Sensor Networks*. China Communications, 8:1:58–64, January 2011.
- [33] J. Chen, X. J. Wu, P. Z. Wen, F. Ye, and J. W. Liu. *A New Distributed Localization Algorithm for ZigBee Wireless Networks*. In CCDC 2009: Proceedings of the 21st Chinese Control and Decision Conference, pages 4451–4456, Guilin, China, June 17–19 2009.
- [34] Y. Kong, Y. Kwon, and G. Park. *Practical Robust Localization over Obstructed Interferences in Wireless Sensor Networks*. In Digest of Technical Papers International Conference on Consumer Electronics, pages 1–2, Las Vegas, NV, USA, January 2009.

4

2D-localization

Both this chapter and the next one focus on localization algorithms. This chapter concentrates on two-dimensional positioning. The position is calculated in the (two-dimensional) plane of the anchors. As outlined in the previous chapter, special care is taken to keep the underlying physical aspects. Therefore, this chapter starts with a simple and accurate statistical linear regression procedure to calibrate the propagation model (equation 3.3) of the w-ilab.t test bed. Next, the accuracy of this model is used to select “well behaving” anchors. The previous chapter further revealed that preprocessing is needed to improve the localization. Our preprocessing uses a min-max procedure on the distances and extends it with an elimination of bad measurements based on the accuracy of the model. Next, the preprocessing executes a maximum likelihood on the distances to further improve the accuracy. For the actual conversion of distances to a position, two different algorithms are presented: the first algorithm LiReFLoA (Linear Regression based Fast Localization Algorithm) is based on a linear regression model and the accuracy of this model. For the second algorithm LiReCoFuL (Linear Regression based Cost Function for Localization), a new cost function is developed and uses this function in a maximum likelihood algorithm. With experiments, we will show that the position errors of both LiReFLoA and LiReCoFuL are smaller than these of the MMSE localization algorithm.

4.1 Selection of anchors

A test was set up using the w-ilab.t test bed on the third floor, The nodes with numbers 10, 19, 32, 39, 40, 42, 44, 48 and 50 were down for maintenance during our experiment (see figure 2.1 (a)). There are 47 active nodes. Every node broadcasts 240 packets. Every packet consists of 100 bytes, the inter packet delay is 25ms. This brief recovery time between packets allows devices to prepare for reception of the next packet. Transmission is at channel 26 in order to avoid Wi-Fi interference (see section 2.3.3). Upon swapping sending nodes, the test bed is idle for 3.5s. The test is performed at a transmit power level of 0 dBm. Every second, available data was sent to our log file.

The test bed only counts valid RSSI-measurements: our software detects corrupted and lost packets; these packets are excluded in the averaging process. Microwave ovens with a traditional power supply only radiate at the positive peaks of the mains supply (see section 2.3.3.5). Our software is able to distinguish valid and invalid packets. Indeed, it uses the RSSI_VALID bit (which is generated by the CC2420 chip [1]) and it averages only over valid RSSI-measurements. Therefore, RSSI-measurements from packets sent at the negative peaks of the main supply are recorded correctly. Fast fading fluctuations can be averaged out by considering a large number of RSSI-reading levels.

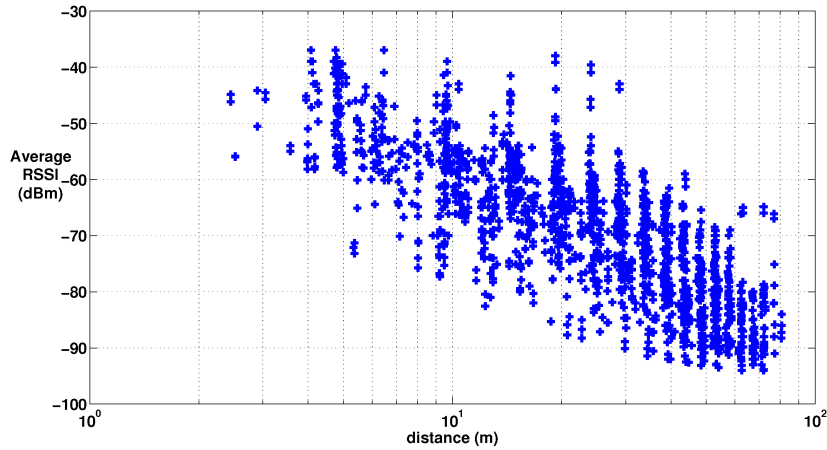


Figure 4.1: Scatter plot of all reported averaged RSSI-readings in function of the distance (on a logarithmic scale)

A first look at figure 4.1, where RSSI versus the distance (expressed in m) between all sending and receiving nodes of our building is plotted on a semi-logarithmic scale, confirms the dominance of multipath fading in indoor environments. Basically the graph consists of 1942 RSSI-logarithmic distance pairs,

where the distance is expressed in meter. The 47 nodes broadcast 240 packets to all other nodes and the average of the RSSI reported by the receiver and its distance to the sender gives one point. Packets below the sensitivity level of a receiver are not reported and thus not presented in the graph. For example at a distance of 24 m RSSI-values between -40 dBm and -84 dBm are encountered. Alternatively a RSSI of -65 dBm corresponds to distances ranging between 5.5 m to 77 m which actually covers almost the whole building. This large RSSI variability is also found in other experimental studies in industrial indoor environments [2]. It is obvious that in such a realistic environment physical relationships can not be applied as such. We therefore use standard statistical tools to solve this problem. We assume a pre-existing sensor network with a large number of nodes, which is a realistic scenario for future dynamic wireless indoor environments. In a first step we select well behaving anchor nodes [3] from all active nodes and calibrate them to their individual propagation parameters according to their underlying physical behavior.

4.1.1 Selection of anchors based on the linearity of their calibrated path loss model

Figure 4.1 looks like a wide monotonously decreasing stripe clipped by a horizontal line at the bottom because of the sensitivity of the receiver. This suggests there still is a linear relationship between RSSI and the logarithmic distance. As already discussed in section 3.2.1, this relation defines a first and very important step in range-based RSSI localization algorithms. Therefore, we further transform equation 3.3 to:

$$\text{RSSI(dBm)} = \text{intercept} + \text{slope} \times \log_{10}(d) \quad (4.1)$$

This equation 4.1 will be intensively used throughout this book, because its (graphical) interpretation is straightforward. Indeed, *intercept* denotes the RSSI measured at a distance of 1 m, where the $\log_{10}(d)$ equals zero. Likewise, *slope* represents the sign-sensitive gradient of the RSSI-logarithmic distance line. Hence, for the theoretical approach in equation 3.2 we instantaneously get an intercept point of -40.2 and a slope of -20 for distances smaller than 8 m. For distances greater than 8 m the standard proposes a steeper slope (-33) and thus proposes a model based on the two ray ground model. Please note these values are invalid for sending power levels different from 0 dBm and frequencies different from 2.4 GHz. Please also note that for Wi-Fi the RSSI is not calibrated in dBm [4] and an extra step is needed to include the relationship between RSSI and dBm. In this section we will calibrate the test bed by determining the intercept point and slope with the linear regression technique, like in [5]. Basically, the regression technique is a least squares problem [6]. Therefore, the Moore-Penrose pseudo-inverse matrix (see section 3.3.2) is used to find the regression line (intercept and slope) out of the

RSSI-measurements and the (known) distances, like in [7]. This is done for each sender (and all other 46 nodes receiving). The use of a considerable amount of receiver nodes improves the quality of the regression model. Next, the r-squared (RSQ) values are calculated for each sender (and all other nodes receiving) with a linear regression tool: RSSI in function of the logarithm of the distance (in m). RSQ is the square of the Pearson correlation coefficient [8]. A zero indicates that there is no linear fit, and with a 1.0 all points lie on a straight line. The Pearson correlation coefficient is positive for ascending regression lines and negative for descending regression lines. For the RSSI at full power the RSQ is varying from 0.89 on node 43 to 0.37 on node 21 in figure 4.2.

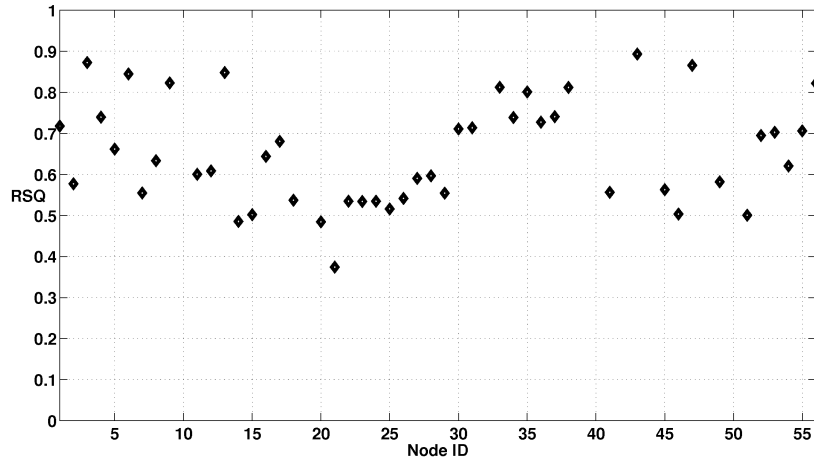


Figure 4.2: RSQ correlation coefficient of RSSI versus logarithmic distance for the sending nodes on the 3rd floor. Some nodes are highly RSSI-logarithmic distance correlated.

The higher the RSQ, the higher the quality of the anchor nodes but also the smaller their quantity. We suggest to choose a RSQ of 0.8 (corresponding to a Pearson correlation coefficient of -0.89) resulting in 10 anchor nodes. The top 10 best fitting nodes are marked with a square in figure 2.1. They are in the extremities of the building, not in the corridors. These nodes are located in positions less sensitive to multipath fading and are considered as more informative ones. This is also confirmed by another empirical study in an indoor environment [9]. In figure 4.3 the intercept points and the slope factors are used to produce graphs of the 10 best fits with $RSQ > 0.80$. A mean of these 10 best fits is also included. This line is constructed using the average of the 10 intercept points and the average of the 10 slopes as a new intercept point and slope respectively.

Figure 4.4 gives the reported RSSI versus the logarithmic distance with the best node 43 sending and all others receiving. Please recall there are 47 active nodes.

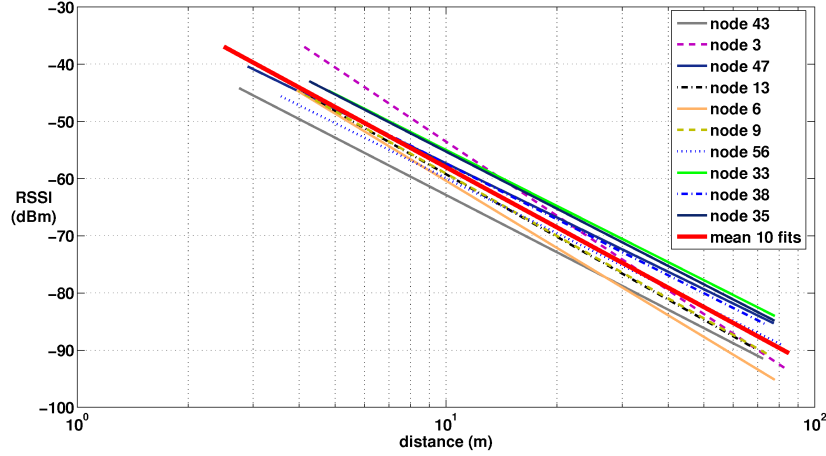


Figure 4.3: RSSI versus logarithmic distance calibration of the 10 best anchors.

Hence, in this half-duplex communication system there are 46 potential receiving nodes. In this figure the measurements from 9 receivers are below the sensitivity level. So there are 46 minus 9 or 37 measurement points. The graph also includes

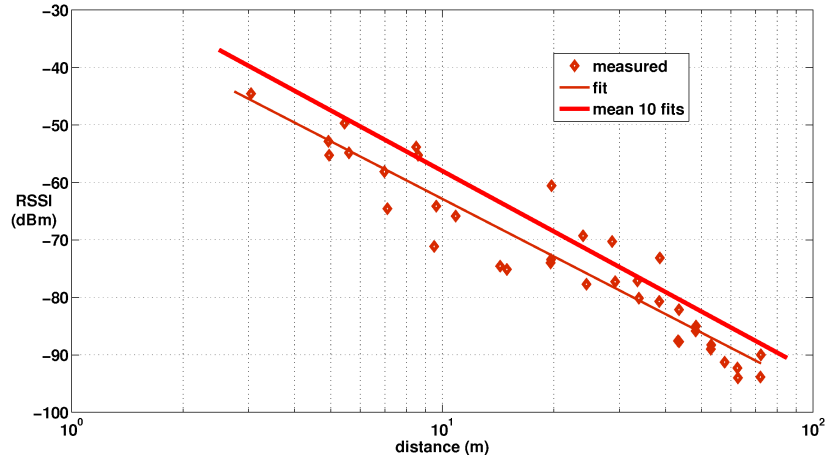


Figure 4.4: RSSI versus logarithmic distance for well behaving sensor node 43. The node has acceptable multipath fading at all receiving nodes.

a linear regression of the measured RSSI and the mean of the 10 best fits. The last two lines are almost coinciding. Measured points above the fit are certainly due to (constructive) multipath fading as explained in section 2.3.2, whereas measured points below the fit are due to (destructive) multipath fading or attenuation.

Figure 4.5 gives the worst performing node 21 (which is in the longest corri-

dor). For this central node the measurements from only one receiver are below the

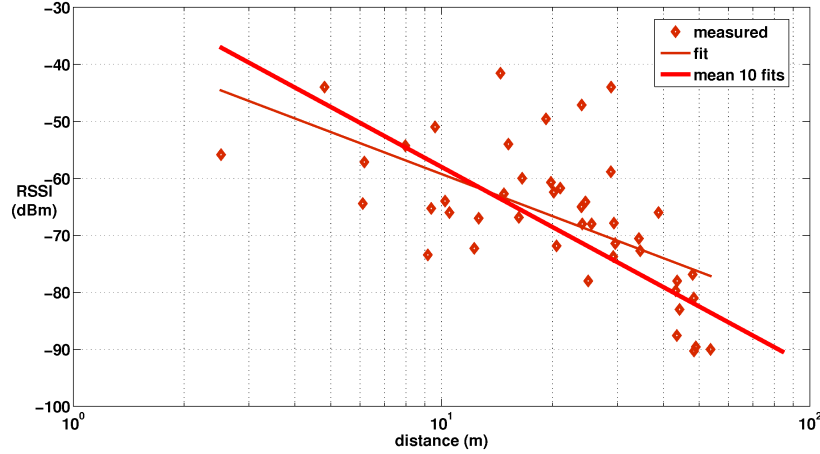


Figure 4.5: RSSI versus logarithmic distance for bad behaving node 21. The slope is less steep.

sensitivity level, yielding 46 minus 1 or 45 measurement points. The fit and the mean of the ten best fits are no longer parallel. The slope of this node is not so steep, multipath fading boosts the tail and the higher attenuation at low distance cuts the head (the reason for this higher attenuation is the passing through concrete to a nearby node). Furthermore, there is a greater difference between the measurements and the fit. Adding more nodes with lower RSQ will pile up with flatter slopes. Calibrating the 10 best sending nodes with their respective intercept point and slope will improve the RSSI to distance conversion. These intercept points and the slopes calculated with known distances can now be used to calculate any distance with a measured RSSI as input. Hashemi [10] reports many researchers added an error correction to equation 3.3 and minimized it in their investigation. In [5, 11, 12] this is e.g. done for the attenuation of walls. Please recall from section 2.2 that many walls of our building are drywall. Passing through them does not affect the RSSI-measurements considerably. Likewise, propagation paths from the sending nodes in the staircases will pass through the same high attenuation concrete material for all the other nodes and therefore mainly influence the intercept point. Because all anchors are individually calibrated, our model absorbs the latter effect. Although a model that includes walls can have a positive effect on the position accuracy, it will add complexity.

4.1.2 Selection of anchors with low standard error

With the 10 best correlated sending nodes, every node in the network hears at least 3 nodes, which should be sufficient for the two-dimensional localization of a subject [13]. Some localization protocols work fine with anchor nodes in the corner [14]. Unfortunately our building (and many others) is not a square and it is obvious that the absolute accuracy of far nodes will be lower. The challenge now is to find a few extra central nodes in order to increase this distance accuracy. Linear regression assumes a constant standard deviation for each point on the regression line [6]. There are two conditions for having a good RSQ. If the slope is flat and the points are spread around this line, a low RSQ will result. On the other hand if the slope is not flat, but the points are spread too far around the line, there will also be a bad correlation. The standard deviation is a measure for how close the points lie around the regression line and is therefore important for determining ranging bounds. We propose an empirical approach for determining and optimizing bounds on the distance. Indeed, any standard linear regression tool can calculate the standard error on the vertical axis which is an estimate of the standard deviation. It can be shown [6] that

$$s^2 = MSE = \frac{\sum_i e_i^2}{n - 2} \quad (4.2)$$

where s is the estimate of the standard deviation, MSE the error mean square (or residual mean square), e_i the residuals (deviation of the observation around its own estimated mean) and n the number of observations. Please note that the division is by $n - 2$ and not by $n - 1$. This comes from the fact that two degrees of freedom are lost: one for calculating the *intercept* and one for calculating the *slope*. This makes s an unbiased estimator of the standard deviation [6]. This means that the expected value of s is equal to the true value of the standard deviation being estimated.

We suggest to swap the RSSI- and the logarithmic distance-axis and to redo the linear regression resulting in the standard error on the logarithm of the distance. We define the *Error_on_Distance* parameter as 2 times the estimated standard deviation of the logarithmic distance frequency distribution. Please note the RSQ is invariant to the axes swap: the RSQ-values don't change when axes are swapped.

The probability density function (pdf) of a normal random variable Y is given by:

$$pdf(Y) = \frac{\exp(-\frac{1}{2}(\frac{Y-\mu}{\sigma})^2)}{\sqrt{2\pi}\sigma} \quad (4.3)$$

where μ and σ are the mean and the standard deviation of the normal random variable Y respectively [6]. Equaling Y to the logarithmic distance $\log_{10}(\frac{distance}{1m})$ and σ to s (the latter only valid for large n , we will come back on this topic later),

we immediately get:

$$pdf(distance_{log}) = \sqrt{\frac{2}{\pi}} \frac{\exp\left(-2\left(\frac{distance_{log} - \mu_{log}}{Error_on_Distance}\right)^2\right)}{Error_on_Distance} \quad (4.4)$$

In figure 4.6 this frequency distribution is plot on a linear distance scale for mean logarithmic distances μ_{log} of 1.3 (20 m), 1.4 (25 m) and 1.7 (50 m) with the same *Error_on_Distance*. It can easily be verified that with this constant *Error_on_Distance* the absolute distance error in meter increases with increasing logarithmic distances. Therefore, central nodes in addition to the 10 anchor nodes

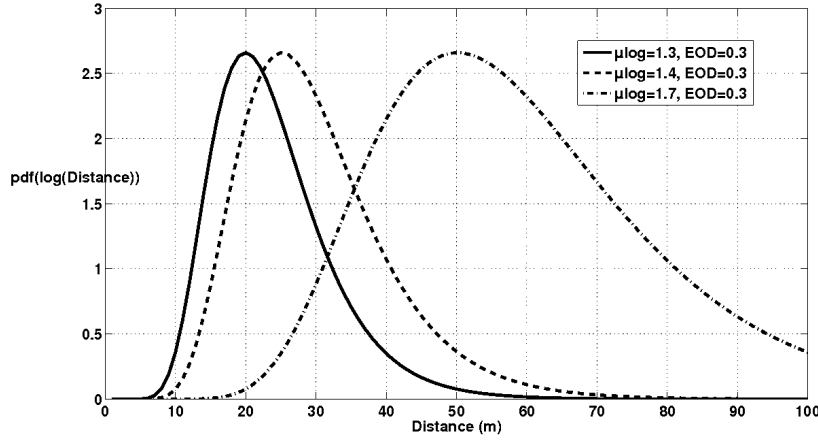


Figure 4.6: Frequency distribution for different points lying on the logarithmic distance versus RSSI regression line.

in the extremities of the building, will improve absolute distance accuracy. RSSI-measurements that are too low (too much attenuation from the walls or too much destructive multipath fading) will result in distances at the right-hand side of the peak of the distribution and will thus correspond to estimated distances that are too high. RSSI-measurements that are too high (certainly too much constructive multipath fading) will result in distances that fall at the left-hand side of the peak of the distribution and will thus correspond to estimated distances that are too low. The above mentioned property is widely used in the design of our positioning algorithm in section 4.3.

Figure 4.7 gives the *Error_on_Distance* for each node. Node 3 is the node with the lowest *Error_on_Distance* (0.21). Please recall from figure 4.2 that this is also the previously selected second best anchor node based on correlation. As can be seen on figure 4.8, where the RSQ is plotted versus the *Error_on_Distance* for each node, the nodes with the lowest errors correspond to the nodes with the highest RSQ, but there are also 2 new nodes in the center: node 30 and node 31.

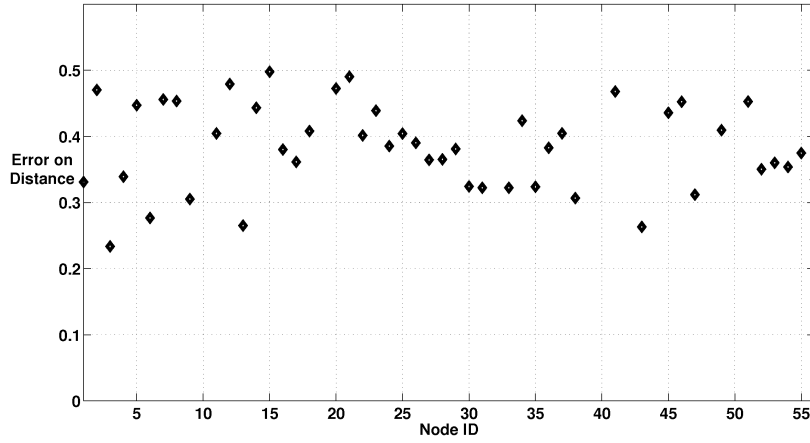


Figure 4.7: *Error_on_Distance* (defined as two times the estimated standard deviation of the logarithmic distance frequency distribution) for the sending nodes on the 3rd floor. Some central nodes have low maximum *Error_on_Distance*.

These nodes are marked with a circle in figure 2.1 (a) and figure 4.8. Including these nodes in our set of good nodes implies a lower absolute error on the position. Measuring a logarithmic distance 1.5 ± 0.3 (31.6 m) will result in a distance of

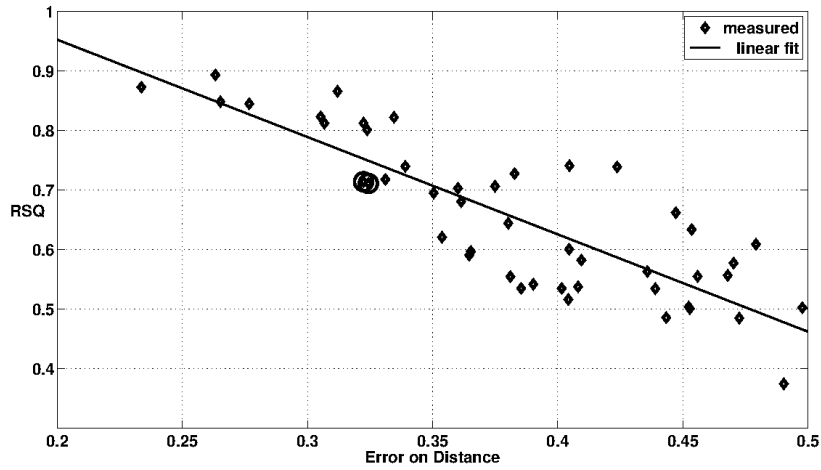


Figure 4.8: RSQ correlation coefficient versus *Error_on_Distance* showing a good linear fit.

$10^{1.2}$ to $10^{1.8}$ or 15.8 m to 63.1 m or 31.6 m (-50% +100%). These error percentages are in a multipath environment an order of magnitude greater than the 10.1% encountered in less complex environments [15], although equally well calibrated.

Please also note that the error percentages in plus and minus are asymmetric, because the distance on error is a logarithmic value and thus the geometric mean needs to be considered.

A typical circle corresponds to the calculated distance, obtained with the measured RSSI and the calculated slope and intercept for the corresponding anchor node.

With the *Error_on_Distance* we produce lower and upper bounds for the expected distances. From now on, we will extensively use “small” and “large” distance circles; by definition these correspond to circles with the radii of these bounds, respectively. The radius of the small distance circle corresponds to the radius of the typical circle divided by $10^{\text{Error_on_Distance}}$. The largest distance circles radii are the typical circles radii multiplied by $10^{\text{Error_on_Distance}}$.

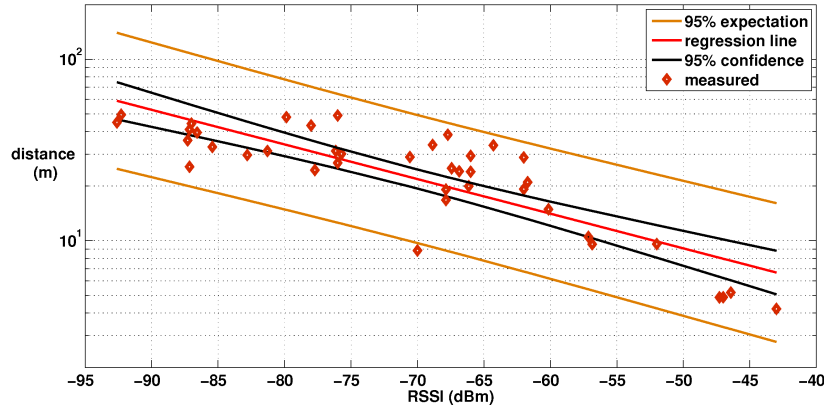


Figure 4.9: Regression line with probability intervals and expectation intervals (also hyperbola) for anchor 30.

In figure 4.9, the regression is shown for node 30 with the swapped axes and two curves are added: the curves with the 95% expectation and 95% confidence intervals. These curves are hyperbola [6] with as center X-coordinate the average of the RSSI regression points and as center Y-coordinate the average of the logarithm of the distance regression points. If an individual logarithmic distance-RSSI experiment is done 95% of the measurements will be found between the expectation lines. The area between the confidence lines (or confidence band) shows where the regression line is expected for the given confidence level. Around the center point the intervals are the smallest. Towards both ends of the regression lines, the intervals become wider. This can be explained by the fact that the uncertainty on the slope increases as the regression point is no longer in the center. Using this property gives an insight in the quality of our regression model: anchor nodes at the extremities of the building have their maximum accuracy at the geometric

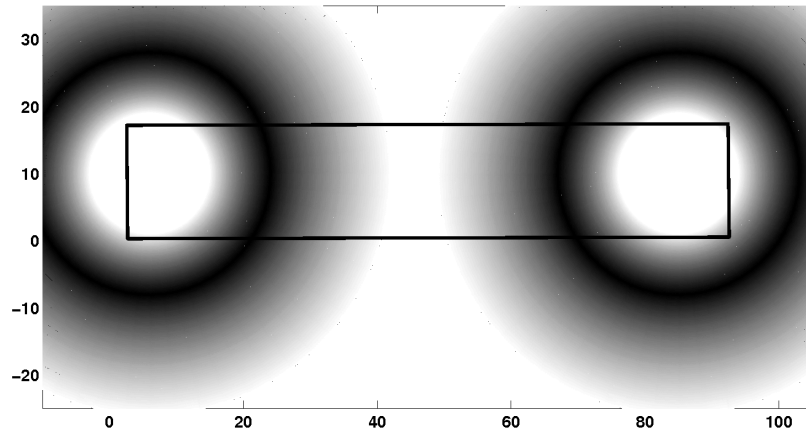


Figure 4.10: Accuracy of the regression model for two anchors placed at the extremities of the (rectangle shaped) building. The darkest intensities correspond to the greatest accuracy. Placing anchors in the center will improve accuracy.

means of the minimum calculated distance (around 4 m) and the maximum calculated distance (around 83 m) or about 18 m. This is illustrated in figure 4.10 with shaded rings for two anchors at the extremities of the building. At the geometric center of these rings the accuracy of the regression model is maximal, because this corresponds to the central (X,Y) coordinates in figure 4.9. Moving away from this central point in this figure results in wider confidence intervals resulting in a lighter gray in figure 4.10. The rectangle represents our building. Rings of central anchors will intersect both other rings and thus have a positive effect on the accuracy. These central anchors unfortunately cannot improve the accuracy of the model at the extremities of the building because there these are also working at the end of their regression lines.

4.1.3 Form factor of the building

This item discusses the shape of a building for a localization algorithm. The upper part of figure 4.11 presents a rectangular building, measuring 10 m x 90 m. The small red circles are the anchors, placed at the extremities of the building, and the blue circles show the estimated distance circles, calculated with the RSSI-measurements of the anchors and the propagation parameters (equation 3.2). The circles form a vertical stripe and the lateral resolution is very low. In the lower part of this table the building is a square with dimensions 30 m x 30 m (or exactly the same surface). The distance circles are now coming from orthogonal directions: the lateral resolution is much better. Most buildings, unfortunately, are not a square, and in the remainder of this chapter we will develop localization algorithms

in a rectangular office environment.

4.1.4 Complexity of the calibration process and robustness against environmental changes

In the previous subsection the best anchors are selected from the active nodes and each anchor is calibrated with its own propagation constants (slope and intercept). This section discusses the complexity of such a calibration process and the robustness of the system against environmental changes. After the experiment on the test bed the measurements are imported in Matlab. For each sending-receiving pair the RSSI is averaged and forms a RSSI-matrix, where element $a_{i,j}$ corresponds to the averaged RSSI reported from receiver j with sending node i . With the

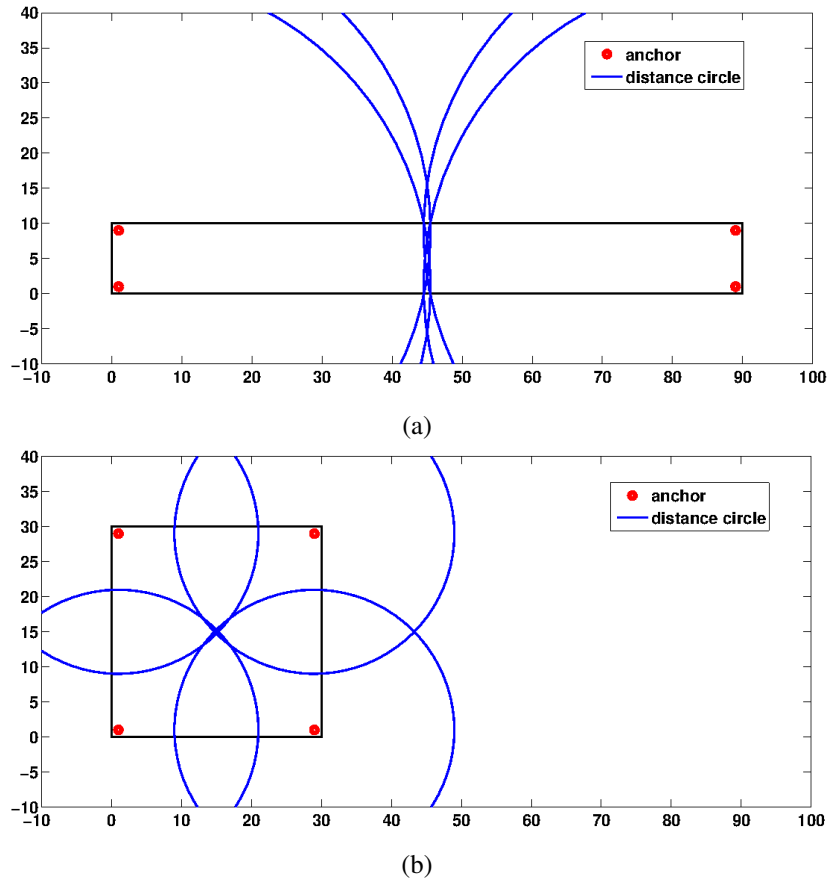


Figure 4.11: Effect of the shape of a building on the longitudinal resolution. (a) rectangular building and (b) square building.

(known) position of the nodes a distance matrix is formed too. Because our nodes are fixed this matrix remains the same for all experiments. On the corresponding rows of both matrices our software now calculates the slope, intercept, RSQ and *Error_on_Distance*. It selects the 10 best correlated nodes and completes the selection with the two first nodes in the sorted *Error_on_Distances* vector, which are not in the best correlated list yet. At this stage the anchors are selected and calibrated. The location information of the remaining nodes is discarded in further procedures and the anchors are used to localize these nodes. Using a single power level and sending 240 packets (with a 25ms inter packet delay) over a single channel the whole calibration phase is done in less than 10 minutes, including the measurements. Calibration can be easily redone, e.g. when a node becomes inactive.

Sending less packets increases the effect of time fading because there are less averaged RSSI-measurements. Sending over more channels and averaging the RSSI-measurements will reduce the effect of different reflections.

For a fingerprinting based algorithm, a training phase is not only typical, but also essential. During this time consuming phase (usually several days) a large database is filled with RSSI-measurements. Next, during the final phase a RSSI-measurement is compared with all records of this database and the record with the best match gives the estimated position. We will not follow this approach in our work: our already deployed nodes are used to quickly and automatically characterize the propagation constants (slope and intercept) of the medium. The knowledge of these two parameters is sufficient to localize (not previously stored) targets.

4.2 Preprocessing

Chapter 3 concluded that preprocessing data before an indoor localization algorithm is tremendously important. Simple mathematical methods, like those discussed in that chapter, are challenging. In presence of outliers, however, the performance of most localization algorithms deteriorates. Our preprocessing detects these outliers in an early stage. Table 4.1 gives a schematic description of the preprocessing algorithm. Each procedure will be explained in the following subsections. In the preprocessing steps we take the underlying physical aspects into account and incorporate simple Min-Max and Maximum likelihood procedures. These concepts, however are not applied to find a position, but to improve the distance measurements. Furthermore, the quality of the regression model (see section 4.1) is used to extend the Min-Max algorithm on the distance to eliminate bad distance measurements. Section 4.3 and section 4.4 will show preprocessing to be effective in reducing the position errors.

Procedure Min-Max

If there are larger than expected large circles for the hardware used:
delete them

Procedure Ignoring_Circles

Delete the typical circles containing at least 5 others if there are no
typical circles contained by at least 5 others

If there are typical circles contained by at least 5 others:
delete them.

Procedure Maximum Likelihood

Group the anchors based on distance

Are there small distance circles contained by other small distance
circles of other group members?

If yes:

new typical circle radius all members=geometric average of
 m calculated distances

$$new_Error_on_Distance = \frac{(Member_Error_on_Distance)}{\sqrt{(\#Remaining_Typical_Circles-1)}}$$

adapt small and large circles according to $new_Error_on_Distance$
continue

If no:

continue

Table 4.1: Schematic description of the preprocessing

4.2.1 Min-max algorithm

It is very important to only feed good input to a localization engine. In this section a resembling technique to the min-max algorithm on the position (see 3.5) is used to detect destructive multipath outliers.

Our software gathers data and produces overviews like table 4.2, table 4.3 and table 4.4. As an example node 4 (table 4.2), node 55 (table 4.3) and node 54 (table 4.4) are considered, but this is done for all targets. In the first column we find the transmitting anchor, the second represents the reported RSSI by the target, the third gives the corresponding calculated distance (taking into account each anchor's own propagation parameters), the fourth the *Error_on_Distance* from section 4.1.2, the fifth the lower and upper bounds for expected distances (radii of the small and large distance circles, see section 4.1.2), calculated from the previous two columns. E.g. for anchor 3 in table 4.2 a calculated distance of 11.3 m with an *Error_on_Distance* of 0.234 gives an expected “small distance” of 11.3 times $10^{-0,234}$ or 6.6 m and an expected “large distance” of 11.3 times $10^{+0,234}$ or 19.4 m. The last column gives the RSQ from section 4.1.1. Table 4.5,

Anchor ID	RSSI (dBm)	Calculated Distance (m)	Error On Distance	Expected Distance (m)	RSQ
3	-55.9	11.3	0.23	6.6 – 19.4	0.87
6	-63.9	12.3	0.28	6.5 – 23.2	0.84
9	-59.3	10.1	0.31	5.0 – 20.4	0.82
13	-62.6	12.3	0.27	6.7 – 22.7	0.85
30	-78.0	35.6	0.32	16.9 – 75.2	0.71
31	-89.8	108.1	0.32	51.5 – 227.2	0.71
33	-90.4	122.2	0.32	58.2 – 256.7	0.81
47	-89.6	105.7	0.31	51.2 – 216.8	0.87

Table 4.2: Data available for positioning node 4

Anchor ID	RSSI (dBm)	Calculated Distance (m)	Error On Distance	Expected Distance (m)	RSQ
13	-89.3	66.9	0.27	36.3 – 123.2	0.85
30	-68.9	20.2	0.32	9.6 – 42.7	0.71
31	-74.0	38.8	0.32	18.5 – 81.5	0.71
33	-44.0	4.6	0.32	2.2 – 9.7	0.81
35	-66.3	21.4	0.33	10.2 – 45.2	0.80
38	-61.9	13.7	0.31	6.8 – 27.8	0.81
43	-62.4	9.7	0.26	5.3 – 17.8	0.89
47	-58.3	10.7	0.31	5.2 – 22.0	0.87
56	-56.3	7.7	0.33	3.6 – 16.6	0.82

Table 4.3: Data available for positioning node 55.

table 4.6 and table 4.7 show the real distance between the anchors to node 4, 55 and 54 respectively.

Section 4.1.1 found 10 good anchors, having a good linear regression between the RSSI reported by all receivers and the logarithmic distances between sender and receiver. Section 4.1.2 added two good central anchors having a low *Error_on_Distance*. Furthermore, it revealed where the anchors are the most accurate. Now, we start polishing up the measurements. Comparing table 4.2 and table 4.5, a large error between calculated (122.2 m) and exact distance (72.0 m) from anchor 33 is seen. Although not an outlier here (the minimal estimate 58.2 m is less than 72.0 m) this measurement can disturb any positioning algorithm. More extreme examples are found with other targets with up to 134 m (1.5 times the size of the building) calculated distance and up to 275 m maximum distance

(three times the size). Multipath fading resulted in flatter slopes of the regression lines and extrapolating these lines causes these large errors.

Anchor ID	RSSI (dBm)	Calculated Distance (m)	Error On Distance	Expected Distance (m)	RSQ
3	-84.1	51.3	0.23	29.9 – 87.8	0.87
6	-85.1	43.0	0.28	22.8 – 81.4	0.84
9	-88.9	66.1	0.31	32.7 – 133.6	0.82
13	-77.1	31.0	0.27	16.9 – 57.1	0.85
30	-70.0	21.7	0.32	10.3 – 45.8	0.71
31	-46.0	6.3	0.32	3.0 – 13.3	0.71
33	-73.4	36.8	0.32	17.5 – 77.3	0.81
35	-60.0	13.9	0.33	6.6 – 29.2	0.80
38	-54.4	8.1	0.31	4.0 – 16.4	0.81
43	-70.0	16.4	0.26	8.9 – 30.0	0.89
47	-73.6	32.8	0.31	16.0 – 67.3	0.87
56	-73.6	26.7	0.33	12.4 – 57.7	0.82

Table 4.4: Data available for positioning node 54

Anchor ID	Real Distance (m)
3	13.0
6	12.2
9	12.6
13	15.4
30	43.1
31	43.4
33	72.0
35	72.9
38	68.2
43	67.3
47	72.2
56	75.8

Table 4.5: Real distance between target node 4 and the 12 anchors.

Like in [16] a combination of a min-max algorithm with a maximum likelihood algorithm is needed to improve these situations. Unlike them we suggest to drop all measurements where the calculated distance is greater than a limit because the

Anchor ID	Real Distance (m)
3	82.5
6	77.8
9	73.9
13	68.4
30	33.7
31	33.9
33	4.8
35	12.9
38	15.4
43	10.9
47	6.8
56	5.0

Table 4.6: Real distance between target node 55 and the 12 anchors.

Anchor ID	Real Distance (m)
3	52.7
6	47.9
9	43.9
13	38.4
30	8.8
31	5.6
33	25.5
35	24.7
38	20.0
43	19.6
47	24.5
56	28.1

Table 4.7: Real distance between target node 54 and the 12 anchors.

correction to that limit can lead to a large error. Based on equation 3.2 a maximum distance for our hardware is calculated. Indeed considering a sending power level of 0 dBm and a receiver sensitivity of -92 dBm, equation 3.2 gives a maximum distance of 82.8 m. RSSI-measurements resulting in higher distances are not in accordance with the hardware used, and are ignored. 35 individual measurements are dropped in this way. For all nodes this additional step improves the distance

error. If a distance calculation is higher than the limits for the hardware, it probably comes from a connection with either too much attenuation or too much destructive multipath fading. In both cases it comes from anchors operating at the end of their regression line. Applying this algorithm on node 4 (in table 4.2) results in the elimination of the RSSI-measurement of anchors 31, 33 and 47.

The blue lines in the floor plan of figure 4.12 show the (direct) transmission paths between node 4 and the anchors 31, 33 and 47. Like in figure 2.1, drywall walls are presented by dark green solid lines, the solid gray lines are concrete wall and the black squares and black circles are the selected anchors. When electromagnetic radiation is obliquely incident on a wall or floor, less power will be transmitted through the wall than would occur at normal incidence [17]. The red pentagrams represent the regions where this situation occurs.

- The higher attenuation in the path between node 4 and anchor 31 is probably due to the scattering at two pentagrams.
- The path between node 4 and anchor 47 crosses a concrete block completely, furthermore there is a region of oblique incidence.
- The path between node 4 and anchor 33 passes 14 scattering nodes and is attenuated by 6 drywall walls.

These paths are not representative for the majority of the transmission paths, and are considered as outliers. Because the slopes for these anchors are flatter than those of the theoretical model (equation 3.2), these outliers can be detected and deleted.

4.2.2 Elimination of circles

In this section, the regression based technique is evaluated on its accuracy. The model accuracy, discussed in section 4.1.2 is helpful to decide whether too small or too large distance circles need to be ignored. Indeed, figure 4.9 and figure 4.10 illustrate that the model accuracy is best in the middle of the regression line. The elimination of circles is based on this knowledge. Table 4.8 illustrates where node 4, 54 and 55 are on the regression lines of the anchors.

Node 54 is representative for a target in the center of the building and node 55 is a good example of a target in an extremity of the building. These nodes perfectly demonstrate the accuracy of the model. With column 3 of table 4.2, table 4.3 and table 4.4 and the exact position of the anchors typical (dot marked) circles can be drawn for locating the nodes. This is done in figure 4.13 for node 4. The radius of these typical circles is the calculated distance (based on RSSI-measurements) to the target, represented by the dark square. If all measurements were on the regression line, all typical circles would intersect in the target. The too

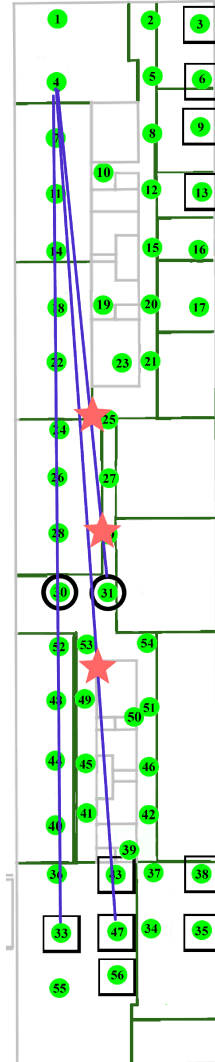


Figure 4.12: The blue lines represent transmission paths between node 4 and anchor 31, node 4 and anchor 47, and also node 4 and 33. These paths are not representative for the majority of the transmission paths. They have high attenuation, due to the passing through walls and the occurrence of oblique incidence on walls.

large (red) typical circles of anchors 31, 33 and 47 contain all other typical circles. The RSSI-measurement is too low and in the logarithmic distance–RSSI graph the

measurement is too far above the linear regression line. As explained in section 4.2.1, these circles have been deleted.

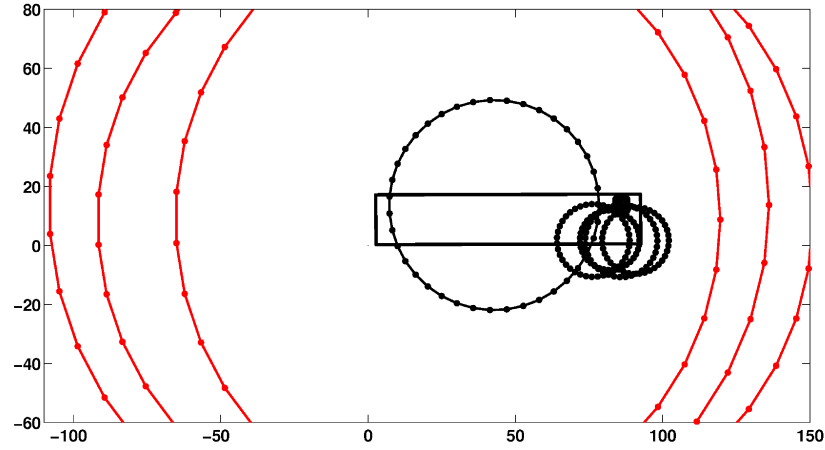


Figure 4.13: Drawing of (dot marked) typical circles for positioning (squared) target 4. The red circles are ignored because they are larger than the limit for the hardware.

	Anchor 3-6-9-13	Anchor 30-31	Anchor 33-35-38- 43-47-56
node 4	(lower) end	(higher) end	(higher) end
node 54	center	(lower) end	center
node 55	(higher) end	(higher) end	(lower) end

Table 4.8: Where are node 4, node 54 and 55 on the regression lines of the anchors?

With two measurements (from two different anchors) of the same target there are four possibilities: they are both too small, they are both (slightly or much) too large, the first one is (slightly or much) too small and the second is (slightly or much) too large or vice versa. Averaging in the first two cases gives no improvement, because the closest distance to the target is corrected in the wrong direction. Averaging in the last two cases only improves if there are no outliers. Deleting bad measurements and outliers is a major preprocessing step in our environment with heavy multipath fading.

Now, we suggest expanding the elimination of bad measurements based on containing-contained circles and the quality of the anchors. Figure 4.14 shows the typical circles for node 55 constructed with table 4.3 and the exact position of the anchors. We observe that the largest (yellow) typical circle (belonging to

anchor 35) contains five others. The smallest (pale blue) circle at the left-hand side (belonging to anchor 33) is contained by 3 others. Node 55 is representative for all nodes at the extremities of the building: there are no circles that are contained by minimum 5 others and there are more than 5 containing circles.

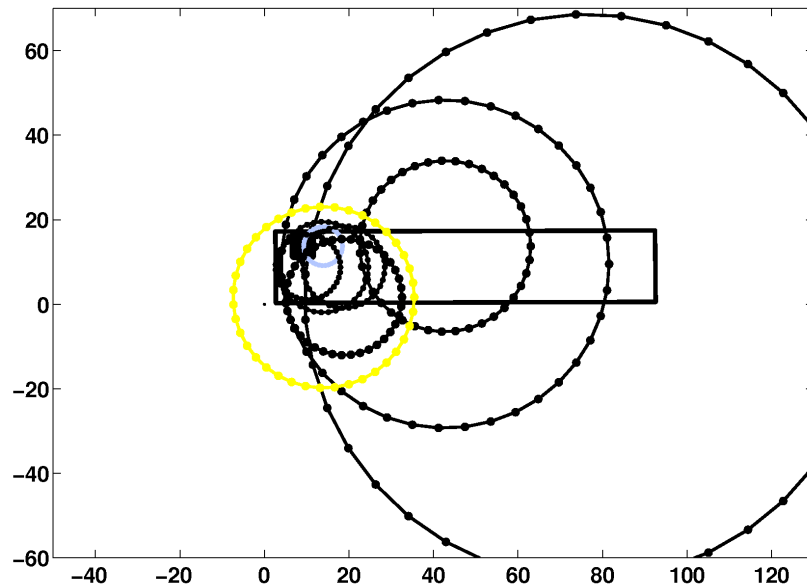


Figure 4.14: Drawing of (dot marked) typical circles for positioning (squared) target 55. All measurements from anchors at the right-hand side and the center of the building have low accuracy for targets at the left-hand side. The largest (yellow) circle at the left-hand side is ignored because it contains too many other typical circles.

For targets at the left-hand side extremity of the building all anchors are operating at the end of their linear regression line. The left-hand sided typical circles are at minimum distance and all others (including the central) at maximum distance of their respective regression line. The accuracy of the smallest typical circles increases when the target shifts to the center, thus these circles are not deleted. Between the large amount of large typical circles there exists not only inaccuracy but also redundancy. In this situation, all typical circles containing at least 5 others will be ignored.

The figure 5 is dependent on the total number of sending anchors (received by the target). When this figure is set too low, also good measurements may be eliminated. This will imply the number of measurements is not large enough for the (good) positioning. When it is set too high, also bad measurements may be allowed. This will have a negative impact on the next step of our preprocessing:

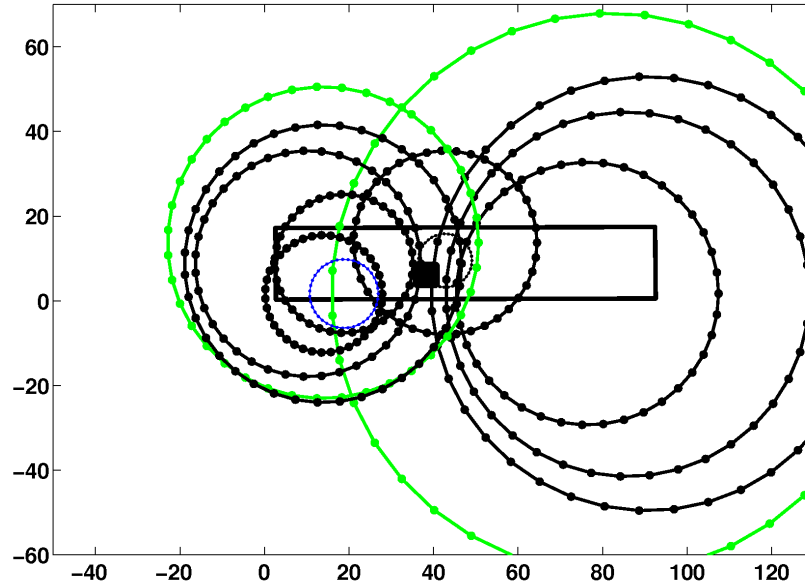


Figure 4.15: Drawing of (dot marked) typical circles for positioning (squared) target 54. The blue circles is deleted because it is contained by more than 5 other circles. The green circles are kept and can be used in the maximum likelihood algorithm on the distance.

the maximum likelihood algorithm.

Likewise, for targets in the center of the building the anchors at the extremities are not extrapolating their regression lines, thus the largest circles shouldn't be deleted now. The anchors in the center of the building are operating at low distances. Too small central typical circles are contained by more typical circles (coming from anchors at both sides of the extremities) than typical circles at the extremities. A representative example is node 54. A closer look at figure 4.15 reveals the green typical circles contain minimum 5 other typical circles and the blue circles are contained by at least 5 other typical circles. Hence, when there are both minimum 5 containing and 5 contained typical circles only the too small (blue) circles are deleted. Please note that the remaining typical circles are more consistent. The too large (green) typical circles will be used in the maximum likelihood algorithm on the distance of the next subsection, where they can correct the distances of their group members.

4.2.3 Grouping Anchors

When anchors are grouped a maximum likelihood algorithm can be used. It is very common to use a maximum likelihood algorithm on the position, as described in section 3.4. In our preprocessing, a reduced complexity maximum likelihood algorithm on the distance is incorporated. Neighboring anchors can be grouped and the most likely distance is estimated. This technique is excellent for a preprocessing procedure because it is fast and achieves a large improvement.

Figure 2.1 (a) shows three groups of anchors: nodes 3, 6, 9 and 13 are in the first group, nodes 30 and 31 in the second and the others in the third. These nodes are neighboring anchors in one extremity of the building, the center and the extremity of the building at the other side. We suggest considering the anchors in one group as coinciding if one “small circle” (with a radius that is equal to the lower bound for the expected distance, see section 4.1.2) is contained in the other(s). Please recall from section 4.1.2 the radii of these small circles are available from the fifth column of table 4.2, table 4.3 and table 4.4, and calculated with the *Error_on_Distance*. Under the assumption that the probability distribution of the logarithmic distance is normally distributed for all group members with the same variance, the likelihood function equals the product of these normal distribution functions of the m (see table 4.1) remaining calculated logarithmic distances in each group. This assumption is not strange: in the previous two procedures, outliers have been deleted. Furthermore, just a few high quality anchors have been selected based on the high RSQ and low EOD (which is an estimate for the standard deviation, and hence the variance). Moreover, figure 4.8 shows a good correlation between the RSQ and EOD. This results in the fact that there is not a large variation in the anchor’s EOD. Partial derivation of this likelihood function [6] proves the most likely logarithmic distance equals the average of the calculated distance of the m group members. Hence the most probable distance can be found out of the (geometric) average of the m (see table 4.1) remaining calculated distances in each group. This geometric average is calculated and applied to all anchors in that group. (Please note the position of the anchors remains unchanged.) If a distance estimation of one anchor is too high, it is compensated by the other group members. Also with distances too low this is the case. By eliminating the outliers before this process, the chance of decreasing the accuracy of the good anchors is minimized. If the small circles don’t contain others, the target is probably nearby and the individual anchors are needed to increase the accuracy.

In this section we suggest to use the standard error of the regression model and divide the *Error_on_Distances* by the square root of the number ($m - 1$) to obtain the corrected *Error_on_Distance* for the group members. We divide by $m - 1$ and not by m because one degree of freedom is lost.

In figure 4.16 the maximum likelihood algorithm on the distance is illustrated. All solid circles are small distance circles. Again, the typical distance circles are

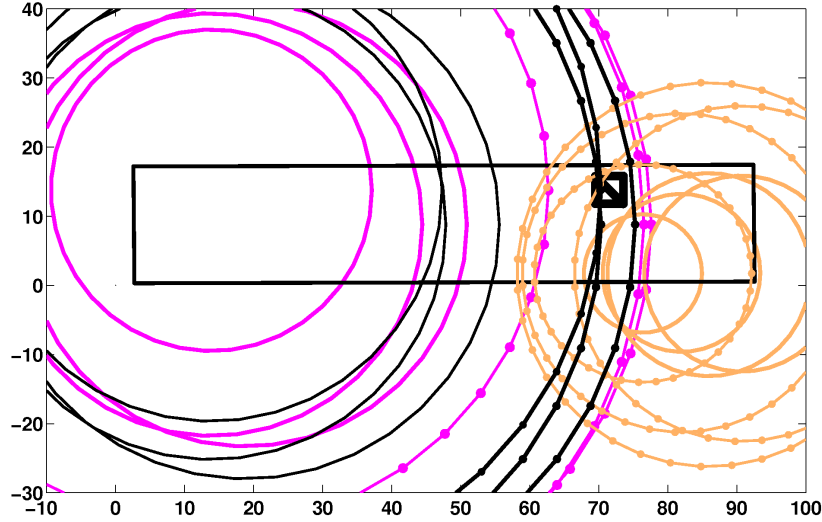


Figure 4.16: Application of the maximum likelihood algorithm on the distance on the magenta group brings the (black dotted) typical circles closer to the target.

dot marked and the target (node 14) is represented by a small square. The orange circles are for the first group with the anchors 3, 6, 9 and 13. Here, none of the small circles is contained by another small circle of any group member. Therefore, the maximum likelihood on the distance is not applied. The magenta circles belong to the third group, consisting of anchors 33, 35, 38, 43, 47 and 56. Please note that the measurements of anchors 35 and 56 are below the noise floor and the measurement of node 38 is above the limits of the hardware (see section 4.2.1). The remaining (magenta dot marked) typical circles are not close to the target. Because the small distance circle of node 33 is contained by the small distance circle of nodes 43 and 47 the maximum likelihood algorithm on the distance is applied. The corresponding radii are calculated using this subsection and result in the black circles. Geometric averaging of 44.7 m (anchor 33), 63.4 m (anchor 43) and 53.5 m (anchor 47) yields 53.4 m as new radius. The magenta distance circles are replaced by the black. Please note the corrected typical circles are now very close to the target. Please also note that the large circles and the circles of a second group consisting of anchors 30 and 31 are not shown in this figure. Including them would have overloaded this drawing.

The (log-normal) probability distributions functions (pdf) for anchors 33, 43 and 47 peak at 44.7 m, 63.6 m and 53.5 m for anchors 33, 43 and 47, respectively, as shown in figure 4.17. These lines are constructed with the measured RSSI at node 14, and the slope, intercept and the EOD of the respective anchors. Just as in figure 4.6, the distributions are wide because of the large EOD (0.32, 0.26, and

0.31, respectively). Multiplication of these distributions results in the maximum likelihood function. The derivative (gradient) of the maximum likelihood function is shown in the gray solid line zero-crossing at a distance of 54.5 m, the most likely distance between the anchors and node 14. This value is very close to the geometric average (53.4 m) of the anchor-target distances, justifying the previously made assumptions of normality and equal variance.

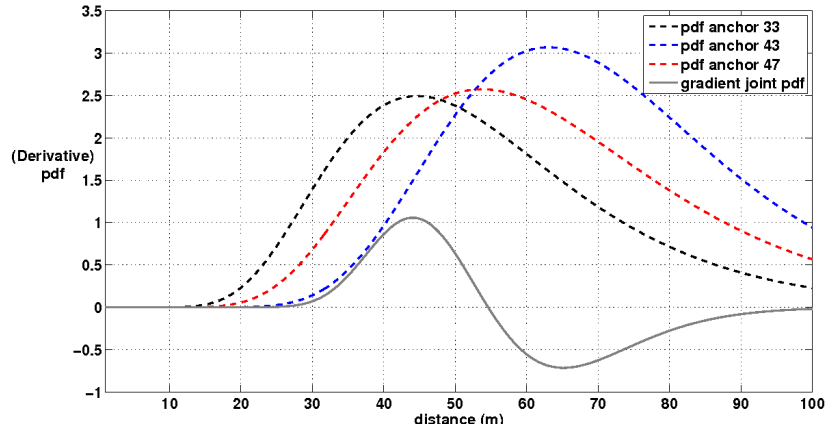


Figure 4.17: Log-normal probability functions for node 14. The gray solid line represents the gradient of the maximum likelihood function. The geometric average (53.4 m) of the abscissa where the respective probability functions peak is a good approximation of the abscissa of the peak of the maximum likelihood function (54.5 m).

4.2.4 Preprocessing results: the distance error

Preprocessing aims at improving the distance ranges before the real positioning. Therefore, this section evaluates the performance of each successive preprocessing step. A good metric for this goal is the distance error. This distance error considers one of the 1942 sender/receiver pairs at a time. The error between the real and the calculated distance results in one point in the cumulative distribution function (cdf) plot. Section 4.3 will show preprocessing to be effective in reducing the position error.

Figure 4.18 gives the cdf of the distance error on the preprocessing steps. The dotted line represents the error distribution for the initial step. The RSSI-measurements of the test described in section 4.1 at full power are converted to distances using the propagation parameters taken from the theoretical linear regression model derived from the attenuation in the IEEE802.15.4 standard when evaluated between 2.5 m and 83 m (see section 4.1.1).

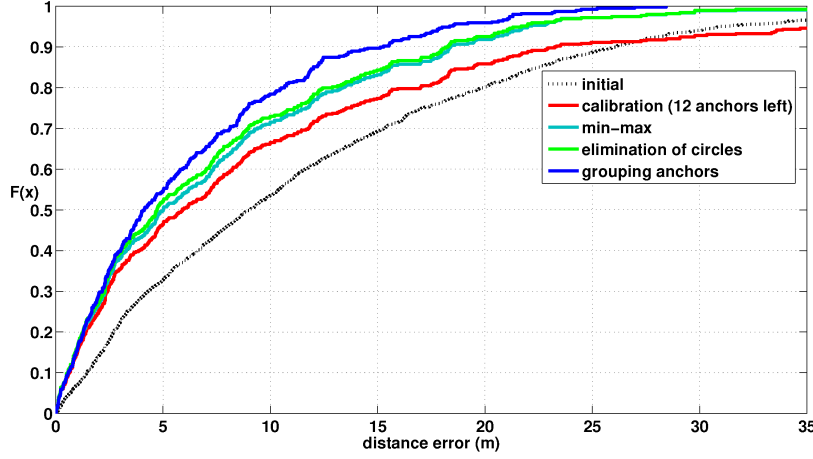


Figure 4.18: Cumulative distribution function plots for the preprocessing steps. Each subsequent preprocessing step further reduces the median of the distance error.

The difference between these distances and the real distance between the corresponding sender/receiver pair defines the (one-dimensional) distance error. The other curves present the cdf for the distance errors after subsequently applying section 4.1.2 (calibration 12 nodes left), 4.2.1 (min-max algorithm), 4.2.2 (elimination of circles) and 4.2.3 (grouping anchors) respectively. For low and medium percentiles, the first preprocessing step already gives an improvement. As can be seen at the upper right side of the figure the calibration of anchors has a negative effect for percentiles above 90. Figure 4.18 also reveals this effect is eliminated when the too large circles are rejected. The subsequent steps of the preprocessing further improve the accuracy.

Please note that all preprocessing steps reduce the median of the distance error from 8.95 m to 4.03 m.

One might argue that there is no significant difference between the “elimination of circles” and “min-max” cdf plots and conclude that the min-max procedure is useless. This conclusion is wrong. Please recall from section 4.1.2 there are 46 potential receivers for each anchor transmission. Therefore, the “calibration” cdf plot consist of (at most) 46×12 or 552 measurements. In our experiment the “min-max” cdf plot consist of 425 measurements. The “min-max” procedure further eliminates only 40 out of these 425 measurements. The other distance errors remain unchanged. The small amount of affected measurements is not sufficient to show a large difference in figure 4.18 for the “elimination of circles”-curve and the “min-max”-curve, but please note that the “elimination of circles”-curve stays above the “min-max”-curve. Examination of the 40 deleted measurements further reveals that the median of the distance error for this subgroup of measurements

equals 8.47 m, or almost the median of the unprocessed distance errors. This subgroup therefore contains outliers. The next preprocessing step is a maximum likelihood algorithm. Please recall from section 3.4 that elimination of outliers is extremely needed for this kind of algorithm.

4.3 LiReFLoA

With the preprocessing of the previous sections good input is achieved for positioning: good anchors are selected, based on high RSQ and low *Error_on_Distance*; bad measurements are eliminated because both too small and too large distance circles are rejected. Furthermore, grouping anchors allows a maximum likelihood algorithm on the distance. All of this results in more consistent distance circles. Our *Error_on_Distance* approach results in small and large distance circles forming rings (washers) with accurate bounds on the distance. These distance rings can be used in an area-based localization algorithm like [18]. In that more theoretical study these authors find the intersection of the rings and find the smallest enclosing circle covering this intersection of rings. Please note that our anchor selection and calibration algorithm provides a method to empirically obtain these rings. The thickness of the rings is not only anchor dependent (see table 4.2, table 4.3 and table 4.4), but (because of the logarithmic property of the *Error_on_Distance*) also distance dependent (see figure 4.6). For this approach, however it is necessary that the rings are not disjunct. This is why the bounds of section 4.2.3 must not be set too conservatively.

Indeed, decreasing the thickness of the rings will increase the risk of these rings to be disjunct. For a few cases this disjunct-ring condition is met in our multipath environment. Instead of increasing the bounds, a low complexity point-based algorithm better suits our empirical data. We present the flowchart in figure 4.19. Mainly it consists of 6 procedures and two decisions. The numbers represent the subsections where the respective items are discussed. Usually the following path 4.3.1, 4.3.2, 4.3.3, 4.3.4, 4.3.5, 4.3.7 is used. In very few cases procedure 4.3.6 is encountered and the use of procedure 4.3.8 is even rarer.

First, the underlying principles of LiReFLoA are discussed.

Section 2.3 revealed that multipath fading is a very annoying problem for indoor localization. Moreover, chapter 3 showed that simple localization algorithms are extremely vulnerable to multipath faded outliers. Indeed, with constructive multipath fading, the RSSI-measurements are (much) too high, and hence the distances too small. With destructive multipath fading, the RSSI-measurements are too low resulting in too large distances. LiReFLoA is based on this physical phenomenon to calculate the position. The first decision (section 4.3.2) is a test on the large distance circles (defined in section 4.1.2). Hence, if the most restrictive large circles don't intersect, there is constructive multipath fading. The second decision

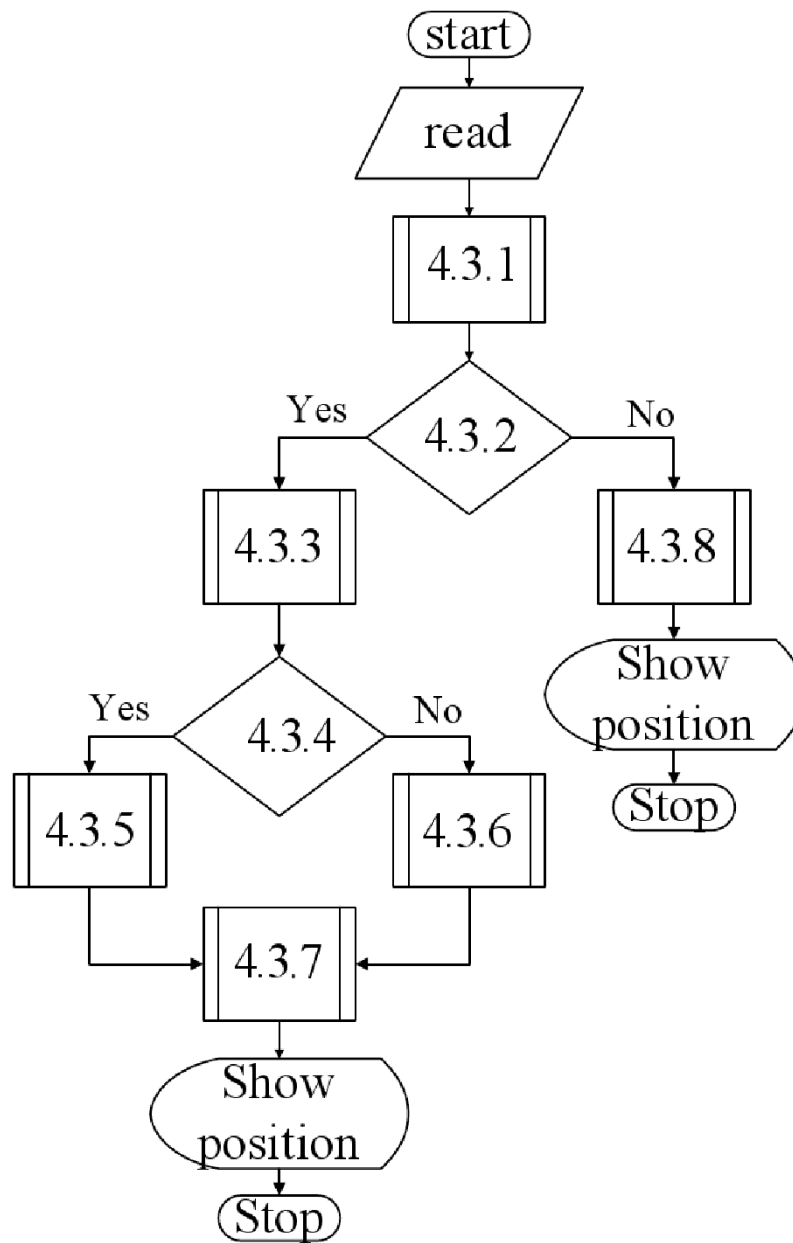


Figure 4.19: Flowchart of the LiReFLoA positioning algorithm.

(see section 4.3.4) is a test on the amount of intersections of typical circles within the building. When none or very few intersections lie within the building, it is likely that there is constructive multipath fading.

4.3.1 Calculate intersection points of the longitudinal borderlines of the building and the large circles

For each target the previous section resulted in small, typical and large distance circles around the anchors. This starting procedure looks for the common intersection of these large circles and the longitudinal borderlines of the building, see figure 4.20.

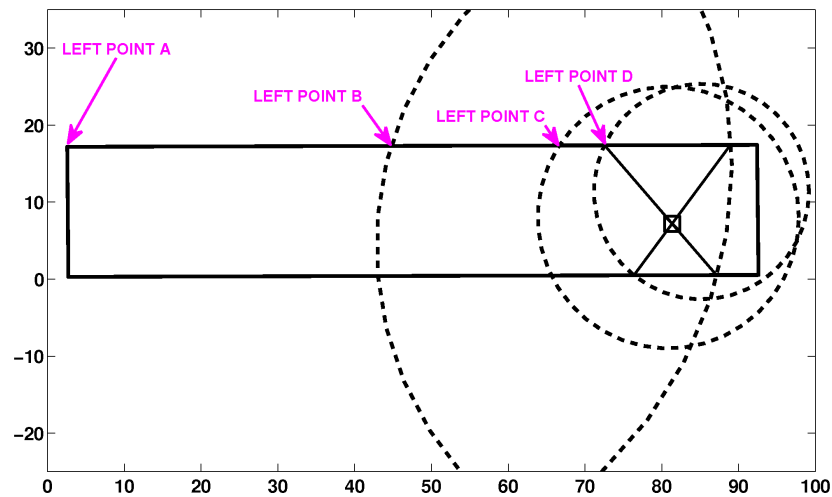


Figure 4.20: Common intersection of the large distance circles and the longitudinal borderlines of the building. Initial position is calculated as the intersection of the diagonals of the trapeze (when initial position cannot be calculated from typical circles).

This is a principle figure and all underlying information as target number, anchor ID ... is left out, in an effort to avoid overloading the drawing. This approach is valid for the remainder of this section. The rectangle represents the building and the large circles are in dashed lines. The intersection points of the large circles and the longitudinal borderlines of the building are calculated. If the leftmost intersection is to the right of the corresponding leftmost corner of the building, it replaces this corner point. Now, the next large circle is considered and its leftmost intersection point replaces the previous if it is more to the right. In this way the left points are changed from point A through D. The interpretation of this point is straightforward: if there are no outliers, all large circles represent effective max-

imum bounds on the distance and the target can not be at the right of the most restrictive left point. This is done for both longitudinal borderlines of the building. An analogous procedure gives the leftmost right points of a trapeze. Without outliers on the constructive multipath fading the large circles will not be too small and the vast majority of the targets will continue on the main track towards 4.3.3.

4.3.2 Is the amount of constructive multipath fading acceptable?

As discussed in section 2.3.2 constructive multipath fading results in too large RSSI-measurements and hence too small distance circles. If the most restrictive large circles do not intersect constructive multipath fading is present. See figure 4.21. The final right point is now situated left to the final left point. The trapeze now has negative sides.

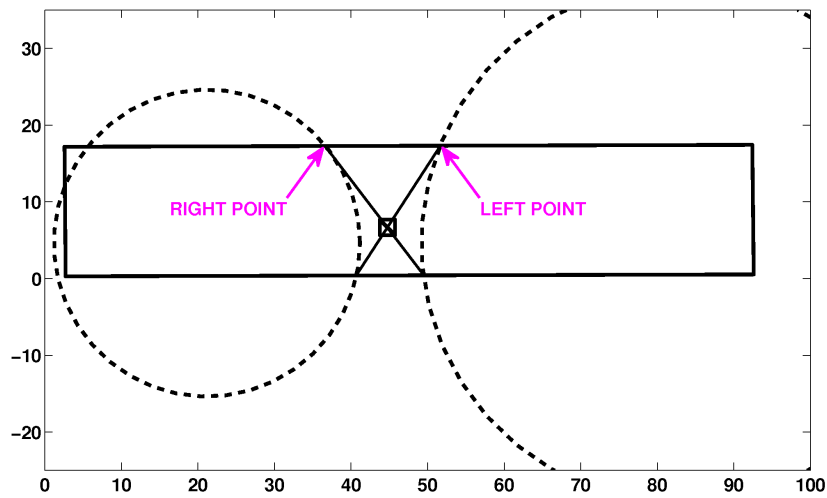


Figure 4.21: Large distance circles in case of constructive multipath fading. Too small large circles result in trapezes with negative sides. Initial position is calculated as the intersection of the diagonals of the trapeze.

When no outliers on the constructive multipath fading are present like in figure 4.20, the trapeze has positive sides. Hence, this is a fast method to decide on the amount of constructive multipath fading. On the positive trapeze sides outcome the algorithm continues with 4.3.3 and with negative sides 4.3.8 is the next step.

4.3.3 Calculate typical circles intersection points within the building

Next, the typical circles intersection points within the building are calculated. There are two possibilities: figure 4.22 shows the situation where the intersection points are inside the building and figure 4.23 illustrates an example where typical circles do not intersect or have intersection points outside the building. Our software counts the intersection points within the building of the most intersected typical circle(s). In the next step a new decision needs to be taken.

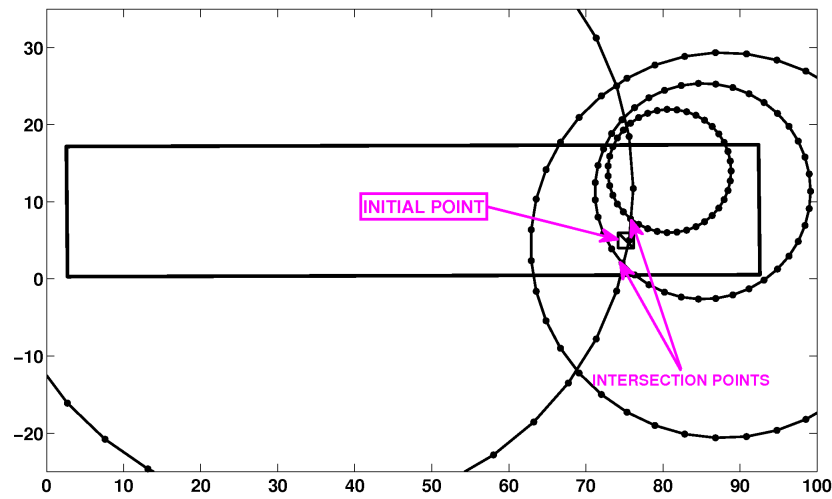


Figure 4.22: Initial position based on averaging of typical circles intersection points.

4.3.4 Is there more than one intersection point within the building?

Preprocessing steps resulted in more consistent distance circles. As discussed in section 4.1.1 constructive multipath results in too small circles. When all typical circles don't intersect at all within the building or when there is only one intersection point of all intersecting typical circles within the building, uncorrected constructive multipath fading is expected. Therefore, the number of intersection points within the building is a good metric for the presence of constructive multipath fading. When there is more than one intersection point within the building, procedure 4.3.5 is executed. If this is not the case, procedure 4.3.6 follows. The difference is a selection of a different initial point. Very few targets follow the 4.3.6 procedure and the vast majority of targets will continue with procedure 4.3.5.

4.3.5 Calculate initial position as the centroid of all typical circles intersection points

A logical next step is the calculation of an initial position. It is the result of a multilateration process, calculated as the centroid of all typical circles intersection points with the typical circle that is intersected most. Figure 4.22 gives an example where the lowest circle is intersected twice. The little square in the middle of the drawing represents the initial position. In a few cases a typical circle is intersected up to 14 times by other typical circles.

4.3.6 Calculate the initial position as the intersection of the diagonals of the trapeze

In figure 4.23 the leftmost typical circle is smaller than the corresponding circle of the same anchor in figure 4.22. It is likely there is more constructive multipath fading for figure 4.23. The typical circles do not intersect within the building, either because there are no intersections or because the intersections are not within the building. Please also note that if the multipath fading for the two rightmost anchors decreases (and these circles become larger) there will again be intersections within the building. Therefore, constructive multipath is expected when there are no intersections within the building. Furthermore, when there is at most one intersection point of all typical circles within the building, the initial point cannot be calculated with the multilateration step of section 4.3.5. In this case the initial position is estimated as the intersection point of the diagonal of the (positive) sided trapeze, calculated from the large circles as in figure 4.21. The procedure described in subsection 4.3.7 will further increase accuracy.

4.3.7 Calculate the estimated position of the nearest intersection points of the longitudinal line and the typical circles

Around the initial point accurate typical circles are expected, because the preprocessing resulted in consistent circles. Please recall from section 3.3 and figure 3.1 that all typical circles would intersect in the target if there were no multipath fading. Furthermore, figure 4.9 shows that the 95 percent confidence intervals (in m) are widest at high distances. Looking in the neighborhood of the initial point at distances larger than the limits of the confidence interval will include outliers. Therefore, we impose an absolute limit on the search region around this initial point. When set too high also non detected outliers will spoil the good measurements and when set too low also good measurements from far anchors will be eliminated. As a limit we propose the same total confidence interval of 22 m for all anchors. Models that are more complex can adapt this margin to the considered

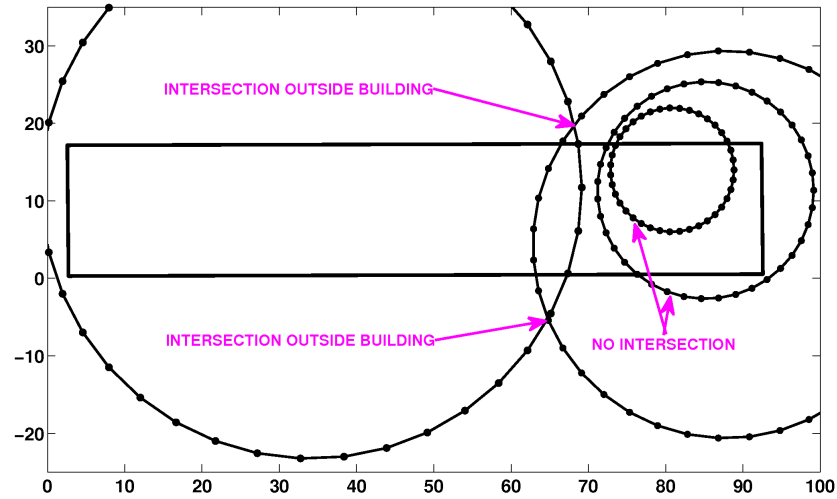


Figure 4.23: Large distance circles in case of constructive multipath fading. Too small large circles result in trapezes with negative sides. Initial position is calculated as the intersection of the diagonals of the trapeze.

anchor. Figure 4.9 shows that for node 30 a 95 percent confidence intervals of 22 m is reached at a distance of 50 m.

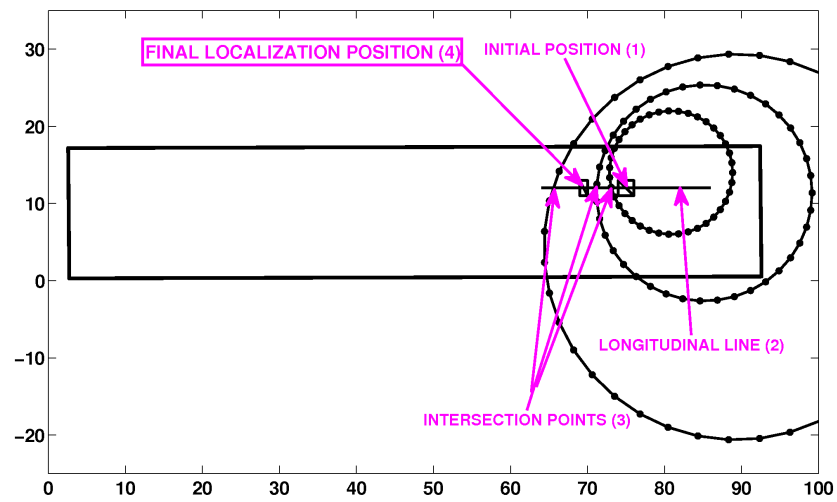


Figure 4.24: Final step of the localization algorithm. A weighted averaging of typical circle points in the vicinity of the initial position results in the final position.

Figure 4.24 represents the final step in the localization. Around the initial posi-

tion (obtained with procedure 4.3.6), a line of ± 11 m is drawn in the longitudinal direction. Please note the limits of the line are symmetrical, while the confidence intervals are not. Models that are more complex could keep these limits asymmetrical. Now, for each typical circle the intersections with this line are calculated. The intersecting point that is closest to the initial point is kept. Multilateration of these kept intersections results in the estimated position. Due to the form factor of the building multilateration in the lateral direction of the building is not performed.

4.3.8 Calculate the position as the intersection of the diagonals of the trapeze

This procedure is executed when the most restrictive large circles do not intersect. Therefore, at least one of the concerned anchor-target paths exhibits too much constructive multipath fading. In our building the preprocessing step reduced the number of affected targets from 5 to 3. Instead of finding out which circle is too small, we allow negative sided trapezes. Considering both circles equally likely to be too small and that the negative sides are small, the position is estimated as the intersection of the diagonals of the trapeze. In figure 4.21 this position is marked with a black square. It is obvious that also the typical circles will be much too small. Hence, searching in the vicinity of this black square (as in section 4.3.5) will not increase the accuracy.

4.3.9 LiReFLoA results: the positioning error

The previous section 4.2 ended with a discussion on the distance error. In this section the distance calculations $\hat{d}_{i,j}$ from the different anchors are grouped and result in one of the 47 target positions at a time. A good metric for evaluation of the complete algorithm is the position error. Here, all distances from the different anchors are grouped, resulting in one of the 47 target positions at a time. Please note that lower distance errors result in lower position errors and that in most algorithms, the position error is higher than the corresponding distance errors. To prove the concept of our method, we use the unselected nodes (non-anchors) as localization targets. However, our technique can also be used for the localization of other targets, for which no location information is available.

For comparison, a more conventional maximum likelihood on the position with a mean square error (MMSE) cost function is implemented. Please recall from section 3.4 the MMSE cost function is given by equation 3.15. For a better understanding of the grid method, the dependency of the calculated distance $\hat{d}_{i,j}$ is explicitly given in equation 4.5

$$(\hat{x}, \hat{y}) = \underset{(x,y)}{\operatorname{argmin}} \sum_{j \in \operatorname{anchor}(i)} (\hat{d}_{i,j}(\text{RSSI}) - d_{i,j})^2 \quad (4.5)$$

Our building is gridded with an intergrid size of 0.50 m in each direction. Next, the exact distances $d_{i,j}$ between the grid points j and the (known) position of anchors i are calculated. Now, the RSSI-measurements at the anchors are converted to the calculated $\hat{d}_{i,j}$ (RSSI) distances. For each anchor, the exact distance to each grid point is subtracted from the calculated distance and this difference is squared. According to equation 4.5, this difference calculation is redone for all contributing anchors. For each grid point, the corresponding squared differences are added. The grid point with the lowest sum of the squared differences is the most likely location.

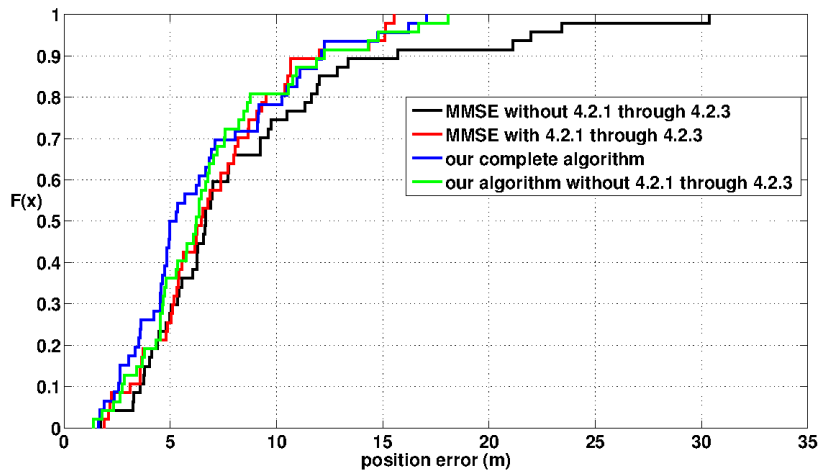


Figure 4.25: Cumulative distribution function plots for the error on the position. Comparison is made between our algorithm and conventional maximum likelihood methods, presented by the mean square cost function (MMSE).

In figure 4.25 the error on the position is shown in our sparse anchor density (12 anchors on a surface of 1512 m^2 or approximately 0.008 anchors per square meter) environment. As a common reference the end of section 4.1.2 is chosen: the best anchors are selected and calibrated. The effect of the presence of the preprocessing steps (section 4.2.1 through section 4.2.3) is studied on both our algorithm and the more conventional maximum likelihood on the position. The figure shows there is an improvement of the preprocessing on our complete algorithm, the median of the error is reduced from 6.22 m to 5.29 m. Furthermore, the worst result is for MMSE without preprocessing steps section 4.2.1 through section 4.2.3. Using our min-max algorithm, eliminating circles and grouping anchors has a positive effect on the position error of the MMSE algorithm. It reduces the median from 6.66 m to 6.45 m. Our complete algorithm has a lower median than the MMSE maximum likelihood algorithm with all our preprocessing steps and it has the lowest high percentiles.

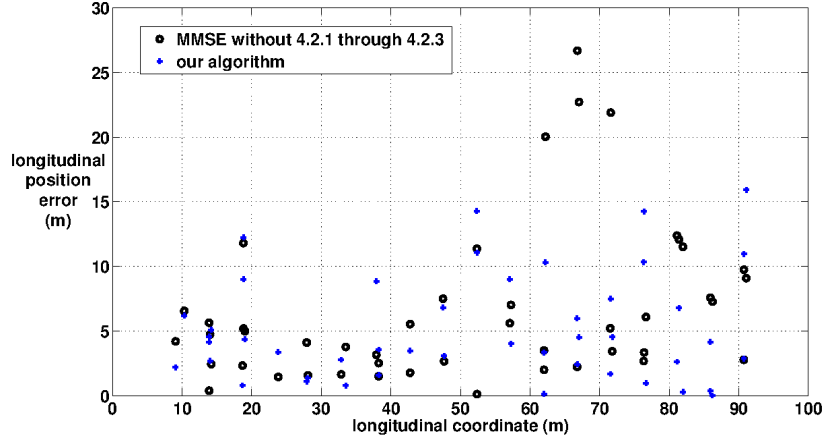


Figure 4.26: Error on the position along the longitudinal coordinate of the building.

The largest errors for both algorithms are made in the longitudinal axis of the building along the longest side of the rectangle. Figure 4.26 shows this longitudinal error for both algorithms and illustrates the existence of many large errors for the maximum likelihood algorithm on the position. Its highest error belongs to node 2 (at a longitudinal coordinate of 91 m), situated not only in an extremity of the building (where the selected anchors are less accurate, see figure 4.10 but also in the longest corridor (with the most constructive multipath fading, see section 2.3.2). Figure 4.27 shows the cumulative distribution function plots for the lateral error on the position. The upper graph is for our complete algorithm and the lower graph is for the maximum likelihood on the position (without preprocessing step section 4.2.1 through section 4.2.3). Our tests show that our algorithm tends to favor the center in a narrow rectangle, while a maximum likelihood on the position rather locates the targets on the longitudinal borders for such a geometry. Indeed, the longitudinal coordinate of LiReFLoA is determined by either the diagonals of trapeze construction (procedure 4.3.6 and procedure 4.3.8), or the averaging of typical circles intersection points 4.3.5. Both procedures favor the center.

Traditional maximum likelihood algorithms can be compared with a bed sheet, having a ball in it. This ball represents the most likely position. At the position of the anchors, this sheet is pulled up with a force proportional to the RSSI-measurement. If all RSSI-measurements are correct, the ball finds the position of the target. This comparison makes clear what happens in narrow rectangular environments: if the anchor's "forces" are not accurate, the ball falls by either one of the waysides.

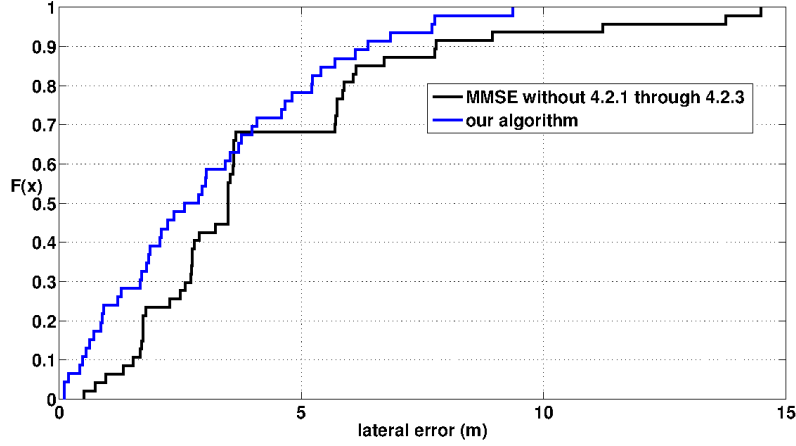


Figure 4.27: Cumulative distribution plot of the lateral error on the position shows our algorithm performs better with a median around 2.6 m (from 3.49).

4.4 Synergy of the preprocessing and LiReFLoA algorithms

In the previous three sections (section 4.1 through section 4.3) we have presented a new approach for localization in a realistic indoor environment, where multipath fading is highly present. Our measurements show that the preprocessing steps decrease the median of the distance error from 8.95 m to 4.03 m. Furthermore, applying our complete algorithm eliminates the outliers and obtains a median of the position error of 5.29 m. A maximum likelihood algorithm with a mean square error cost function has a position error median of 6.66 m. When our preprocessing is applied, not only this median, but also the high percentiles of this algorithm are improved. In our pre-existing sensor network with a large number of sensor nodes, the best available nodes are selected as anchors and calibrated. The whole process can be automated using standard linear regression tools. Time consuming manual fitting and complex fingerprinting is avoided, making it possible to do real-time localization in future dynamic wireless indoor environments.

4.5 LiReCoFuL

Researchers have already been investing a lot of effort in localization-aware applications [19]. Within the DEUS-project [20], a next generation network and service has been implemented by the use of T-mote Sky modules in an elderly surveillance localization system. Modern widely accepted methods use statistics like Bayesian estimators [21–23] and maximum likelihood estimation [24] to improve the ac-

curacy of the position. The maximum likelihood concept has been introduced in section 3.4, and the widely used Minimum Mean Square Error function (MMSE) maximum likelihood estimator for localization has been discussed. Furthermore, in section 4.1, we presented a statistics based automated method to select, optimize and calibrate anchors before offering the RSSI-measurements to our preprocessing algorithm. This preprocessing algorithm, presented in section 4.2, further uses the accuracy of the underlying statistical model to eliminate bad measurements. In section 4.3, we presented a Linear Regression based Fast Localization Algorithm (LiReFLoA) [25]. This tool calculates the position. Here, at the end of this chapter, we use the same selection and calibration method, present a new maximum likelihood cost function and compare it with cost functions that are more traditional, like (MMSE) [26], Relative Location Estimation (RLE) [18] and Reduced Bias Relative Location Estimation (RBRLE) [18].

4.5.1 The LiReCoFuL cost function

The starting point of a maximum likelihood algorithm is a cost function. Several cost functions exist: the simplest and widely used cost function is the minimum mean square error function, presented in equation 3.15. Please recall that $d_{i,j}$ is the Euclidean distance between a point j and an anchor i . Furthermore, $\hat{d}_{i,j}$ denotes the estimated (most likely) distance and is the estimated distance between point j in the x-y plane and anchor i . Although we estimate this distance with the Received Signal Strength Indicator (RSSI)-values of the radio chip and the propagation constants, the log-normal relationship between RSSI and distance is not a prior assumption. Thus, this cost function means that the most likely position is a point in the x-y plane where the sum of squared position errors between estimated and Euclidean distances to the anchors is minimal. Equation 3.15 does not take into account that the underlying physics dictates the relationship between the RSSI and the distance to be semi-logarithmic [27]. Therefore, Patwari et al. start with this assumption and propose the Relative Location Estimation (RLE) cost function [18]:

$$(\hat{x}, \hat{y}) = \underset{(x,y)}{\operatorname{argmin}} \sum_{j \in \operatorname{anchor}(i)} \ln^2 \left(\frac{\hat{d}_{i,j}^2}{d_{i,j}^2} \right) \quad (4.6)$$

where \ln stands for the natural logarithmic function. This cost function implies that the most likely position is a point in the x-y plane where the sum of squared logarithms of the squared quotient of the Euclidean and estimated distance is minimal. Since this cost function is biased (this means that the mean of the estimated position does not equal the Euclidean distance), the same authors suggest a better

cost function with reduced bias (RBRLE)

$$(\hat{x}, \hat{y}) = \underset{(x,y)}{\operatorname{argmin}} \sum_{j \in \operatorname{anchor}(i)} \ln^2 \left(\frac{\hat{d}_{i,j}^2}{C^2 d_{i,j}^2} \right) \quad (4.7)$$

where C is calculated with the propagation parameters and the standard deviation on the RSSI. This standard deviation is estimated with the Cramer-Rao lower bound (CRLB). The authors of [18] notice that $C \approx 1.2$ for typical channels. Therefore, we use this value in this section.

LiReCoFul is based on linear regression and probabilities around a point on the regression line.

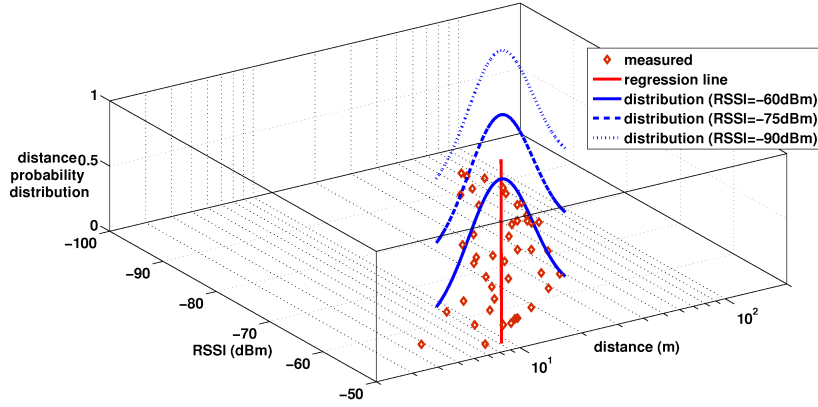


Figure 4.28: Linear regression and distance probability distribution.

The RSSI-distance plane in figure 4.28 presents the measurements for a well behaving anchor with the regression line. This line reduces the mean squared errors, thus the measurements are close to this regression line. A point further away from this line will therefore result in a lower probability of occurrence. The distance probability distribution is shown for three different values of the RSSI in the third dimension of this plot, according to the basics of the linear regression technique [6]. This is also valid for other RSSI-values and thus a kind of tunnel is formed around this regression line. An assumption of linear regression theory is that the y-coordinate values are normally distributed with the same standard deviation. Therefore, the width of the tunnel remains constant for a specific regression line. Having defined an *Error_on_Distance* [25] we were the first to assume a normal distribution on the (logarithmic) distance. Many other authors, including [18], assumed a normal distribution on the RSSI. Because the variables are linearly correlated, both assumptions are equivalent. Our approach however is

more direct, because it outputs distances rather than RSSIs.

$$\log_{10}\left(\frac{\hat{d}_{i,j}}{d_{i,j}}\right) \sim \mathcal{N}(0, \sigma_i) \quad (4.8)$$

Consequently, equation 4.8 is a normal distribution with zero mean and unknown standard deviation. Dividing equation 4.8 by this standard deviation results in a standard normal distribution. For each anchor, the exact standard deviation is estimated from the measurements using the regression technique.

Dividing equation 4.8 by the standard (logarithmic distance) error (or half the *Error_on_Distance* [25]) results in a conversion of this standard normal distribution into a t-distribution [6]. The most likely location is now found by maximizing our cost function:

$$(\hat{x}, \hat{y}) = \underset{(x,y)}{\operatorname{argmax}} \prod_{j \in \operatorname{anchor}(i)} \operatorname{tpdf} \left[\frac{\log_{10}\left(\frac{\hat{d}_{i,j}}{d_{i,j}}\right)}{SE_{\operatorname{anchor}(i)}}, n(i) \right] \quad (4.9)$$

where $\operatorname{tpdf}(t, n)$ denotes the Student's t probability distribution function with n degrees of freedom at the t -value of t [6]. The anchor dependent degrees of freedom $n(i)$ can also be obtained by linear regression: for each sending anchor, $n(i)$ is two units less than the number of receivers with a RSSI-measurement above the noise floor. Indeed two degrees of freedom are lost: one for calculating the mean and one for calculating the standard deviation [6]. When the Euclidean distance of a point in the x - y plane to a particular target equals the estimated distance, the t -value is zero and the t -distribution peaks. This is the case for all anchors. Assuming that the anchors are independent, the overall probability is found by multiplying the probabilities of the individual anchors. Therefore, multiplication needs to be done for all points that are anchors and the cost function needs to be maximized.

Mostly the conjugate gradient algorithm is used to find the extremes of the cost functions (equations 3.15, 4.6, 4.7 and 4.9) [28]. A drawback of this method is that it does not always converge to the wanted extreme of the function, or that it converges to a local extreme [29]. Some authors [30] therefore use this algorithm in combination with another coarse positioning algorithm. In this section we put a grid on our building and calculate the cost function for each grid point. This algorithm is safer because it always finds the true extreme and allows easy visualization, at the expense of extra computational cost.

4.5.2 Theoretical comparison LiReCoFuL and RLE

This section compares the LiReFLoA and RLE cost functions theoretically. An overview is given in table 4.9.

Item	RLE	LiReCoFuL
Propagation model	used in RSSI	used in distance
Dependent variable	RSSI	Distance (axis swap)
Assumption	$RSSI(i, j) \sim \mathcal{N}(RSSI, \sigma_i)$	$\log_{10}(\frac{\hat{d}_{i,j}}{d_{i,j}}) \sim \mathcal{N}(0, \sigma_i)$
Calculate density		
function of	RSSI	$\log_{10}(\frac{\hat{d}_{i,j}}{d_{i,j}})$
with	area under tail normal distribution	σ_i peak Student-t (measurements)
Calculation	multiplication	multiplication
joint density		
Find MLE by	differentiation	keeping multiplication

Table 4.9: Theoretical comparison LiReCoFuL and Relative Location Estimation

Please recall from section 4.5.1 that both cost functions start from the propagation model (equation 3.3). LiReCoFuL swaps the axes. Therefore, the distance (instead of the RSSI) becomes the dependent variable. A next logical step is the assumption that this (statistical) variable is normally distributed. Now, the density function (for one anchor) is calculated. For this density function, RLE uses the (minimum) tail area under the normal distribution. In LiReCoFuL, the (peak) of a Student-t distribution is used. Next, both cost functions assume independence of the anchors and multiply the respective density functions to obtain a joint density function. RLE uses an additional step and differentiates the joint density function, transforming the multiplication into a sum. LiReCoFuL keeps the multiplication, because a multiplication has a smooth gradient in the neighborhood of the targets. A more detailed description of the RLE cost function can be found in [31].

4.5.3 Test conditions for grid based maximum likelihood

This section describes the test conditions for this grid based maximum likelihood algorithms. The iMinds w-ilab.t test bed is used in our experiment. More about this test bed can be found in section 2.2. The second floor is used in this section. In section 4.5.7 the third floor is examined. On figure 2.1(b), this second floor is shown with the position of the 58 nodes. Figure 4.29 shows the position of the active nodes, the anchors and a central target for the LiReCoFuL test on the second floor. The floor is rectangle shaped, but in the center of the floor, there are also outside walls, almost cutting the floor into two smaller rectangles. The floor is gridded with a grid size of 0.25 m in each (lateral and longitudinal) direction. In this section, we use not only the same selection method of best anchors but also

the same calibration method as in section 4.1. Again, each node broadcasts 240 packets of 100 bytes with an interpacket delay of 25ms. Furthermore, transmission is at channel 26 to avoid Wi-Fi interference and upon swapping the nodes, the test bed is idle for at least 3.5s. Our test bed collects the data, we swap the RSSI and (logarithmic) distance axes, perform a linear regression and use regression properties to obtain the well behaving and calibrated anchors. These anchors are marked with a black square in figure 4.29.

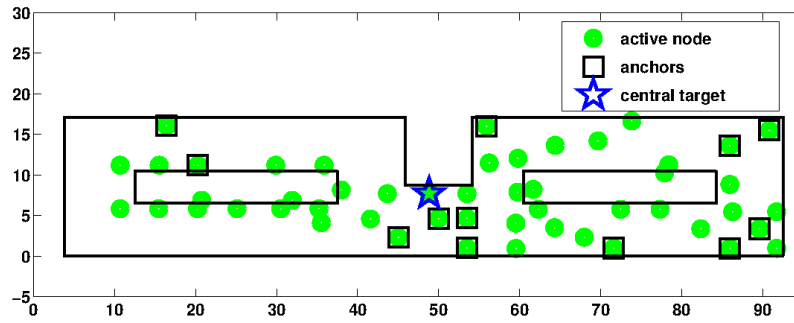


Figure 4.29: Position of the active nodes, the anchors and a central target for the LiReCo-FuL test on the second floor.

4.5.4 Graphical comparison of the cost functions

Now, the different cost functions can be compared. This section initiates this comparison with the plots of the cost functions and the next section follows with a cumulative distribution function plot of the position error.

Figure 4.30 plots the cost functions (equations 3.15, 4.6, 4.7 and 4.9) on a 0.25 m grid (for the same central target) respectively. In figure 4.29 this target is marked with a blue pentagram. This target is chosen randomly. Other targets have similar graphs. For RBRLE, a C -value of 1.2 is chosen. Please recall that MMSE, RLE and RBRLE need to be minimized. For this central target, the Euclidean distances to the extremities of the building are large in the cost function of equation 3.15. This results in the shape of the upper left MMSE graph in figure 4.30.

Near an anchor the denominator of the \ln -argument of equation 4.6 and equation 4.7 is very small. When this point is not the target, the nominator of the \ln -argument is not small. This results in peaks of these cost functions at the anchor locations, forcing the estimated position to the lower values in both the upper right-hand side RLE- and lower left RBRLE graphs of figure 4.30. A large value of the C -value will increase these peaks more pronouncedly. Our cost function for the target can be found in the lower right corner of figure 4.30. Please recall that our cost function needs to be maximized. It has a large gradient around the maxi-

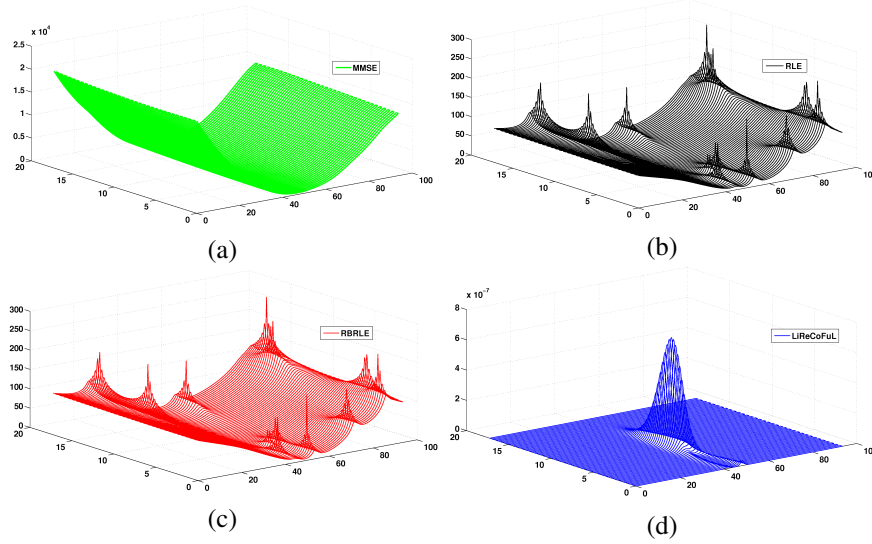


Figure 4.30: Comparison of the different cost functions for a central target (a) MMSE, (b) RLE, (c) RBRLE and (d) LiReCoFuL.

mum. It has less local maxima than other cost functions have local minima. This eases a real-time positioning algorithm based on the conjugate gradient method.

4.5.5 Cumulative distribution function plots of the position error

Our software now calculates the position of each of the 51 active nodes for the different algorithms and compares the results with the exact positions. In figure 4.31, a cumulative distribution function plot (cdf plot) of the position error is given for the different cost functions. The Euclidean distance between the exact and the calculated position presents one position error point in this cdf plot. The MMSE cost function gives the worst results. It has a median of 4.86 m. This can be explained by the fact that the model does not take into account the log-normal relation of the distance and the RSSI. The other medians are 3.23 m, 4.01 m and 3.23 m for the RLE, RBRLE and our cost function respectively.

It can be shown that the frequency distribution of the position error is not a normal distribution. Therefore, non-parametric tests are performed. A Friedman test [32] rejects the null hypothesis that the error distributions are the same for all cost functions. The p-value (defined as the probability that the test statistic is equal to or more extreme than the one observed under the null hypothesis [32]) equals 0.003 or 0.3%. Next, 6 Wilcoxon tests [32] are done, pairwise comparing the position error of the cost functions. E.g. a first test compares the position error for

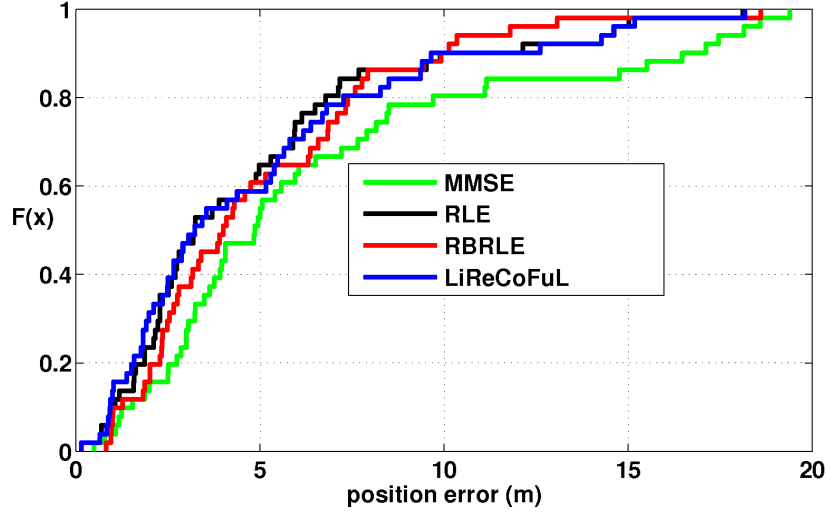


Figure 4.31: Cumulative distribution plot of the position error for the different cost functions on the second floor.

MMSE and LRE (for the same target), a second the error for MMSE and RBLRE, etcetera. These tests confirm that RLE, RBRLE and our cost function result in lower position errors than MMSE. One-tailed p-values are less than 0.05%, 2.9% and 0.05% respectively. The tests further fail to prove a difference between the position error of our cost function and both RLE and RBRLE. Therefore, this subsection shows that the position errors of LiReCoFuL are comparable with those of (RB)RLE and definitely better than those of MMSE.

4.5.6 Execution times

Very fast execution times are needed for real-time localization. This aspect is treated in this section.

At the starting point of this algorithm comparison, anchors are already selected and calibrated. Therefore, those execution times are not treated here. At this stage, a RSSI matrix and a distance matrix are already calculated in Matlab. The $RSSI(i,j)$ matrix consists of averaged RSSI elements reported from receiver j with sending node i . The distance matrix contains elements with the (known) distance between receiver j and sending node i . First, the grid points are calculated. In our 0.25 m gridded building this results in a matrix of 26000 rows and two columns (one for the longitudinal and one for the lateral coordinate). A denser grid will result in a larger matrix and therefore also in larger execution times. Now, the position errors are calculated for each algorithm. On our Dell Latitude D830 position server equipped with Matlab, the average time for calculating one of the 51 positions took

27, 45, 45 and 240 ms for the MMSE, RLE, RBRLE and LiReCoFuL algorithm respectively. Implementing the t-distribution equation [33]:

$$tpdf(t, n) = (n\pi)^{-\frac{1}{2}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \left(1 + \frac{t^2}{n}\right)^{-\left(\frac{n+1}{2}\right)} \quad (4.10)$$

instead of using the tpdf build-in Matlab function will speed up our algorithm. Please note that equation 4.10 is differentiable. This eases the implementation of equation 4.9 in a conjugate gradient algorithm.

4.5.7 The cost function with different scenarios

A similar test as the test described in section 4.5.3 was done on the third floor of the iMinds building. The cdf plots can be found in figure 4.32. The medians of the position errors are 7.05, 4.19, 4.83 and 4.01 m for MMSE, RLE, RBRLE and LiReCoFuL respectively. This confirms the findings of section 4.5.5.

Furthermore, the presence of the longer corridors on the third floor results in higher constructive multipath fading. This explains the fact that the medians on this floor are higher than those on the second floor for the same algorithm.

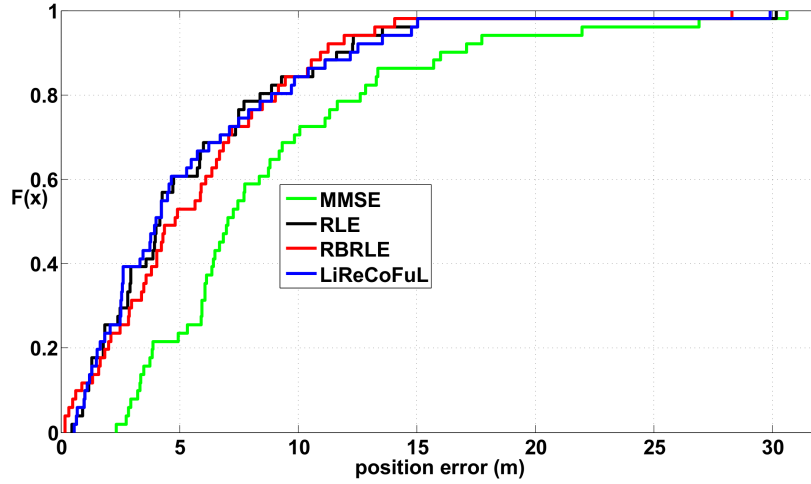


Figure 4.32: Cumulative distribution function plot of the position errors for the different cost functions on the third floor.

4.6 2DLiReFLoA and LiReCoFuL

4.6.1 2DLiReFLoA

The two-dimensional positioning algorithm is a sequential combination of the algorithms described in section 4.1, section 4.2 and section 4.3.

4.6.2 Comparison 2DLiReFLoA and LiReCoFuL

Figure 4.33 shows the cumulative distribution plot of the position errors for the 2DLiReFLoA algorithm (see section 4.3) and the LiReCoFuL algorithm (see section 4.5). The test is done on the third floor, the test conditions are described at the beginning of section 4.1. The anchors are selected and calibrated, as described in section 4.1. The preprocessing of section 4.2 is not applied on the LiReCoFuL algorithm.

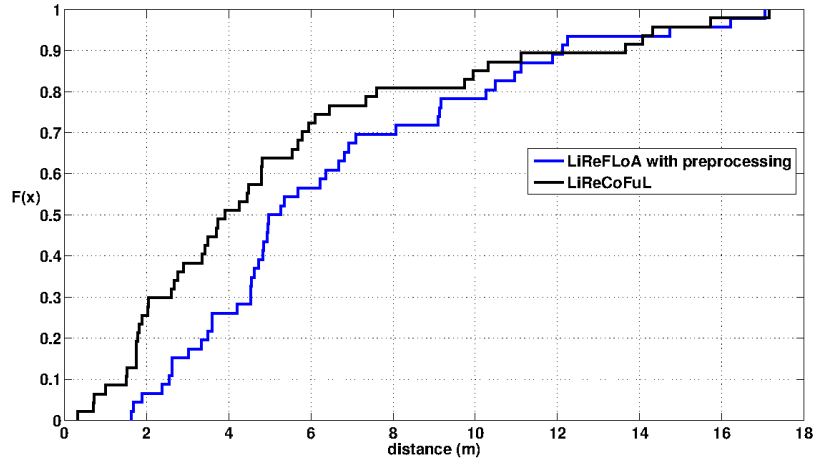


Figure 4.33: Cumulative distribution function plot of the position errors for the LiReFLoA with preprocessing and the LiReCoFuL algorithms on the third floor.

This figure shows that the LiReCoFuL algorithm has a lower position error median (4.26 m) than the LiReFLoA algorithm (5.29 m).

4.7 Conclusions

This chapter concentrated on the empirical development of two-dimensional localization algorithms using the iMinds w-ilab.t test bed. First, a statistical model, based on linear regression has been constructed. This simple model is able to handle large databases and implements the physical aspects of the propagation chan-

nel. In a first step, it has been used to select and calibrate the best available anchors at a certain time. Next, the model accuracy has been investigated in the preprocessing to eliminate bad RSSI-measurements. This preprocessing incorporates a fast maximum likelihood algorithm on the distance. It has been shown empirically that the preprocessing steps reduce the median of the distance error from 8.95 m to 4.03 m in our environment with heavy multipath fading.

Next, two different localization algorithms have been presented. The first algorithm LiReFLoA (Linear Regression based Fast Localization Algorithm) further uses fast geometric principles, based on the model accuracy to obtain a position. The synergy between the preprocessing and LiReFLoA algorithms has been demonstrated: the complete algorithm performs better than the MMSE algorithm. Time consuming manual fitting and complex fingerprinting is avoided, making it possible to do real-time localization in future dynamic wireless indoor environments.

The second new localization algorithm is called LiReCoFuL (Linear Regression based Cost Function Localization) because it uses a linear regression based cost function in a maximum likelihood algorithm. LiReCoFuL not only respects the underlying physics of the propagation model, but also the estimated standard deviation on a sufficient large dataset. It has been shown empirically that the position errors of LiReCoFuL are (just like LiReFLoA) better than a maximum likelihood algorithm with a MMSE cost function. Furthermore, the grid approach reveals that the (RB)RLE cost function has more local minima than LiReCoFuL has local maxima. This enables LiReCoFuL to be used with fewer convergence problems in a conjugate gradient algorithm.

A first test on the third floor shows that LiReCoFuL has a higher accuracy than 2DLiReCoFuL. It is interesting to know if this is also valid in different environments. This is kept as future work.

References

- [1] Texas Instruments. *CC2420 datasheet*. <http://www.ti.com/lit/ds/symlink/cc2420.pdf>.
- [2] A. Azenha, L. Penada, and A. Carvalho. *Error Analysis in Indoors Localization using ZigBee Wireless Networks*. In IECON 2010: Proceedings of the 36th Annual Conference of the IEEE Industrial Electronics Society, pages 2193–2197, Phoenix, AZ, USA, November 7–10 2010.
- [3] S. Alikhani. *Localization in Wireless Sensor Networks*. <http://www.site.uottawa.ca/~casteig/files/csi5140-shafagh-alikhani.ppt>.

- [4] *IEEE Standard for Information Technology – Telecommunications and Information Exchange between Systems – Local and Metropolitan Area Networks – Specific Requirements – Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, 2007. IEEE802.11-2007 Std.
- [5] P. Barsocchi, S. Lenzi, S. Chessa, and G. Giunta. *Virtual Calibration for RSSI-based Indoor Localization with IEEE 802.15.4*. In ICC: Proceedings of the IEEE International Conference on Communications, pages 5–10, Dresden, Germany, June 14-18 2009.
- [6] J. Neter, W. Wasserman, and M. Kutner. *Applied Linear Statistical Models*. MacGraw Hill, New York, USA, 2005.
- [7] Y.-Y. Cheng. *A New Received Signal Strength based Location Estimation Scheme for Wireless Sensor Network*. IEEE Transactions on Consumer Electronics, 55:3:1295–1299, August 2009.
- [8] J. S. Bendat and A. G. Piersol. *Random Data, Analysis and Measurement Procedures*. John Wiley & Sons Academic Press, New York, USA, 1986.
- [9] P. Pivato, L. Fontana, L. Palopoli, and D. Petri. *Experimental assessment of a RSSI-based localization algorithm in indoor environment*. In I2MTC: Proceedings of the IEEE International Instrumentation and Measurement Technology Conference, pages 416–421, Austin, TX, USA, May 3–6 2010.
- [10] H. Hashemi. *The Indoor Radio Propagation Channel*. In Proceedings of the IEEE Vehicular Technology Conference, pages 416–421, Denver, CO, USA, May 1993.
- [11] P. Barsocchi, S. Lenzi, S. Chessa, and G. Giunta. *A Novel Approach to Indoor RSSI Localization by Automatic Calibration of the Wireless Propagation Model*. In Proceedings of the IEEE Vehicular Technology Conference, Barcelona, Spain, April 2009.
- [12] P. Barsocchi, S. Lenzi, S. Chessa, and F. Furfari. *Automatic Virtual Calibration of Range-Based Indoor Localization Systems*. Wireless Communications and Mobile Computing, 12:17:1546–1557, December 2012.
- [13] N. Ha and K. Han. *Positioning Method for Outdoor Systems in Wireless Sensor Networks*. In ISCIS2006: Proceedings of the 21st International Symposium on Computer and Information Sciences, pages 783–792, Istanbul, Turkey, November 2006.

- [14] L. Doherty, K. Pister, and L. El Ghaoui. *Convex Position Estimation in Wireless Sensor Networks*. In Proceedings of IEEE Infocom 2001, pages 1655–1663, Anchorage, AK, USA, April 2001.
- [15] K. Whitehouse and D. Culler. *Calibration as Parameter Estimation in Sensor Networks*. In WSNA2002: Proceedings of ACM International Workshop on Wireless Sensor Networks and Applications, pages 59–67, Atlanta, GA, USA, September 28 2002.
- [16] Y. Kong, Y. Kwon, and G. Park. *Practical Robust Localization over Obstructed Interferences in Wireless Sensor Networks*. In Digest of Technical Papers International Conference on Consumer Electronics, pages 1–2, Las Vegas, NV, USA, January 2009.
- [17] K. Cheung, J. Sau, and R. Murch. *A New Empirical Model for Indoor Propagation Prediction*. IEEE Transactions on Vehicular Technology, 47:3:996–1001, 1998.
- [18] N. Patwari, A. O. Hero, M. Perkins, N. S. Correal, and R. J. O’Dea. *Relative Location Estimation in Wireless Sensor Networks*. IEEE Transactions on Signal Processing, 51:2137–2248, 2003.
- [19] M. Hazas, J. Scott, and J. Krumm. *Location-Aware Computing Comes of Age*. Computer, 37:2:95–97, February 2004.
- [20] DEUS. <http://www.iminds.be/en/research/overview-projects/p/detail/deus>.
- [21] D. Madigan, E. Elnahrawy, R. P. Martin, W.-H. Ju, P. Krishnan, and A. S. Krishnakumar. *Bayesian Indoor Positioning Systems*. In Proceedings of IEEE Infocom, Miami, FL, USA, March 2005.
- [22] G. Chandrasekaran, M. A. Ergin, J. Yang, Y. Chen S. Liu, M. Gruteser, and R. P. Martin. *Empirical Evaluation of the Limits on Localization Using Signal Strength*. In SECON 2009: Proceedings of the 6th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks, pages 333–341, Rome, Italy, June 2009.
- [23] M. Nicoli, C. Morelli, V. Rampa, and U. Spagnolini. *HMM-based Tracking of Moving Terminals in Dense Multipath Indoor Environments*. In EUSIPCO 2005: Proceedings of the EURASIP European Signal Processing Conference, Antalya, Turkey, September 2005.
- [24] T. Roos, P. Myllymaki, and H. Tirri. *A Statistical Modeling Approach to Location Estimation*. IEEE Transactions on Mobile Computing, 1: 1:59–69, January–March 2002.

- [25] F. Vanheel, J. Verhaevert, E. Laermans, I. Moerman, and P. Demeester. *Automated Linear Regression Tools Improve RSSI WSN Localization in Multipath Indoor Environment*. EURASIP Journal on Wireless Communications and Networking 2011, Special Issue: Localization in Mobile Wireless and Sensor Networks, 2011:38, July 2011.
- [26] Y. M. Kwon, K. Mechitoch, S. Sundresh, W. Kim, and G. Agha. *Resilient Localization for Sensor Networks in Outdoor Environments*. In ICDCS 2005: Proceedings of the 25th IEEE International Conference on Distributed Computing Systems, pages 643–652, Columbus, OH, USA, June 2005.
- [27] C. Chang and W. Liao. *Revisiting Relative Location Estimation in Wireless Sensor Networks*. In ICC2009: Proceedings of International Conference on Communication, pages 42–46, Dresden, Germany, June 2009.
- [28] J. L. Nazareth. *Conjugate–gradient Methods*. Wiley Interdisciplinary Reviews: Computational Statistics, 1:348–353, 2009.
- [29] S. M. Kay. *Fundamentals of Statistical Signal Processing, Volume 1: Estimation theory*. Prentice Hall, Upper Saddle River, NJ, USA, 1993.
- [30] Y. Zhao, L. Dong, J. Wang, B. Hu, and Y. Fu. *Implementing Indoor Positioning System via ZigBee Devices*. In Proceedings 42nd Asilomar Conference on Signals, Systems and Computers, pages 1867–1871, Pacific Grove, CA, USA, October 26–29 2008.
- [31] N. Patwari, R. J. O’Dea, and Y. Wang. *Relative Location in Wireless Networks*. In VTC 2001 Spring: Proceedings of the 53th Vehicular Technology Conference, volume 2, pages 1149–1153, Rhodes, Greece, May 6–9 2001.
- [32] J. J. Higgins. *Introduction to Modern nonparametric Statistics*. Duxbury Press, Belmont, CA, USA, 2004.
- [33] Howard W Sams & Co Engineers. *Reference Data for Radio Engineers*. Sams, Indianapolis, IN, USA, 1975.

5

Pseudo-3D localization

Localization in presence of (indoor) multipath fading remains a challenging task. Statistical methods, like statistical multilateration and maximum likelihood estimators are widely used to improve the accuracy of the position. The principle of these algorithms has been outlined in chapter 3. In chapter 4, we presented alternative statistical methods to select and calibrate the anchors, and preprocess the data. Furthermore, two new two-dimensional localization algorithms, Linear Regression based Fast Localization Algorithm (LiReFLoA) and Linear Regression based Cost Function for Localization (LiReCoFuL) have been introduced.

This chapter focuses on three-dimensional localization. However, exact three-dimensional localization is far more complex than two-dimensional localization and usually requires different algorithms and a combination of wireless technologies. Pseudo-3D algorithms use two-dimensional projection techniques to find an object in a three-dimensional space. This reduces the complexity. After the projection two-dimensional localization algorithms are used. In this section, we follow this approach and expand LiReFLoA to obtain a fast pseudo-3D algorithm P3DLiReFLoA with the same number of anchors. Empirically, we will show that execution times barely change.

5.1 Introduction

Table 5.1 gives an overview of what is meant by two-dimensional, three-dimensional and pseudo-3D localization. The classification based on the position of the anchors

and the targets.

Targets to be located in the same plane as the plane of the anchors is called two-dimensional localization. This has been discussed in chapter 4. Three-dimensional localization adds the third dimension: the target is located in the three-dimensional space. This requires that the anchors are deployed in the three dimensions. Sometimes, it is sufficient to find the position of the target in a plane that is different from the plane of the anchors. This is called pseudo-3D localization, treated in this chapter.

Localization	Anchor location	Target location
2D	in one plane	same as anchor plane
3D	in 3 dimensions	in 3 dimensions
Pseudo-3D	in one plane	in one plane not in anchor plane

Table 5.1: Definition of the two-dimensional, three-dimensional and pseudo-3D localization

Our work focuses on experimental RSSI-based WSN indoor localization as in [1–3]. The shortage of experimental results obtained from real indoor test beds, as outlined in [2], is well known. That latter work shows many similarities with ours. Like these authors, we also present a new localization algorithm. The environments, however, are difficult to compare because our test bed is larger (1512m^2 versus 23.2m^2) with the same number of anchors (12 anchors). This results in an anchor density of only 0.008 (versus 0.517) anchors per square meter. Furthermore, our automated calibration method is able to manipulate more measurements. During the selection and calibration method each of the 41 nodes transmits 240 packets to the other nodes. The corresponding RSSI-measurements are reported and averaged. More than 380000 RSSI-measurements are manipulated (somewhat less than $41 \times 40 \times 240$ because not all packets were above the noise floor of the receiver). This is an order of magnitude higher than 12240 RSSI-measurements reported in [2]. Please recall from section 2.3.1 that a higher number of measurements increases the accuracy of the rough measurements, because the fast fading variation is averaged out.

The offline phase of the statistical indoor localization method, described in [4], is based on a Local regrESSion (LOESS) [5] fitting method to build a large RSSI database (called radio map) containing the distribution of the signal strength received at each known location. LOESS divides the independent statistical RSSI-variable in small intervals and performs a regression on these binned data intervals. This offline phase tries to capture the complete distribution of the RSSI-distribution. Next, an online phase involves a maximum likelihood procedure on

the distribution and the measured signal strength. A time-consuming bootstrapping method is used. This method includes resampling the data (typically more than 1000 repetitions are needed): the 95% confidence intervals for the estimated position are obtained by randomly drawing data (with replacement) from a set of data points. Our work is also based on statistics, but takes a completely different approach: the underlying physical (and widely accepted [6, 7]) relationship between the RSSI and the logarithm of the distance (equation 3.2) results in a regression that is faster and simpler, because it is linear in the complete RSSI-variable and does not need data binning. Furthermore our algorithm requires no radio mapping: the knowledge of two parameters (slope and intercept of the regression) is sufficient for the estimation of the position, further reducing the execution time. In this chapter, statistics are also used for comparing results with non-parametric hypothesis testing, where no assumption needs to be made about the distribution of the position error. To our knowledge, this has not been encountered in WSN localization yet. More traditional research uses the cumulative distribution function (cdf) of the position error as well as parametric statistical metrics (mean value, average value and standard deviation) to measure the localization performance [1]. Outliers can affect these parametric parameters substantially, and make the tests and conclusions less reliable.

Very few authors [2, 8] calibrate the propagation parameters to their individual values. In section 4.1 we used linear regression techniques to automate the selection and the individual calibration of the anchors. Here, this chapter uses this technique in a pseudo-3D algorithm.

Three-dimensional indoor positioning is complex and requires a combination of technologies. In [9], a three-dimensional algorithm is presented combining RSSI, time of arrival and sophisticated three-dimensional ray tracing. Ray tracing, which is a widely accepted technique for exact three-dimensional positioning, is based on geometrical optics. It can be applied as an approximate method for estimating the levels of high-frequency electromagnetic fields [10]. With the knowledge of the three-dimensional layout of the building and the materials used, path losses can be predicted. With this path loss the distances can be calculated. This time consuming task can be performed by the use of software tools as in [11]. In this chapter we will not follow this methodology, because it remains tedious.

RADAR-based localization systems [12] and their two-dimensional fingerprinting method, are widely known: in a time consuming training phase a database is filled with RSSI-measurements. In the online phase a measurement is matched with these previously stored measurements. This two-dimensional fingerprinting method can be expanded to the third dimension. In a dynamically changing environment (e.g. changing the position of furniture, presence of persons), however, the time consuming training phase needs to be redone in order to get accurate results [13, 14]. Therefore, [13] proposes an artificial neural network (ANN)

incorporating not only a dynamic fingerprint, but also databases using a linear regression-based tree model mining technique. This approach trades in lack of accuracy for complexity. Effective three-dimensional fingerprinting needs complementary fingerprints of RSSI, temperature, humidity and light [14]. In most cases simple localization algorithms, like Weighted Centroid Location (WCL) [2] are more robust against the variability of the investigated parameters [14].

Other three-dimensional localization systems require a full three-dimensional deployment. At least the double amount is then required: one node on the ceiling and one node on the floor. This solution is mostly used in multi-story buildings, as in [15]. It could be useful in buildings with extremely high ceilings. In these cases the anchors are used efficiently. In most practical situations, however, vertical resolution is not always a primordial matter: e.g. in a museum information system it is more important to know that a person is in front of a particular painting, than the information that he is standing or sitting. Therefore, most localization algorithms don't take the third dimension into account and apply the procedures just as in the two-dimensional localization. In simple RSSI-based algorithms (like LiReFLoA in section 4.3), however, the two-dimensional and three-dimensional propagation paths can differ significantly. When two-dimensional propagation paths are used for calibration in a three-dimensional environment, large errors occur, spoiling the accuracy of the underlying model. Therefore, [16] proposes a complexity-reduced multilateration for three-dimensional localization using super anchors (anchors with pairwise antenna positions whose coordinates only differ in the z-axis). The projection of three-dimensional super anchors reduces the three-dimensional localization to a two-dimensional multilateration algorithm. In this chapter we extend this approach of reduced complexity with very low additional computing time: no super anchors having two antennas are needed, because our three-dimensional calibration is performed with a mobile node, just beneath the two-dimensional selected anchors. Very few extra calibrations are needed: the person to be located makes an initial walk through the building and beneath the anchors, he triggers the mobile node.

5.2 Test conditions for pseudo-3D test

Our hardware consists of the iMinds w-ilab.t test bed. More about this test bed can be found in section 2.2. Only the third floor is used in this chapter. On this 16.8 x 90 x 2.65 m floor, there are 41 active nodes, represented by small green circles in the floor plan of the third floor in figure 5.1.

The nodes are fixed in a plane, approximately 0.15 m below the highly conductive ceiling. An anchor selection results in the best available quality (as discussed in section 4.1) anchors at a given time. These are represented by the black squares. Like in the two-dimensional case (see chapter 4), there is a lot of con-

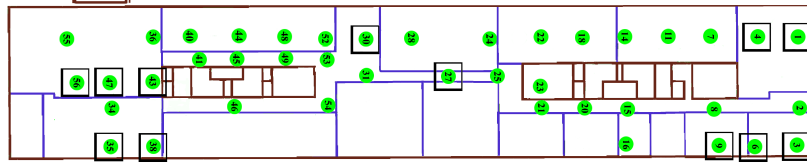


Figure 5.1: Position of sensor nodes on the third floor of the iMinds office building. Drywall walls are presented by blue solid lines. The solid brown lines are concrete wall. The black squares are the selected anchors (see section 4.1).

structive multipath fading, due to the presence of long corridors and the highly conductive ceiling (see section 2.3.2). Because the walls mainly consist of drywall (these walls are represented by blue solid lines), the presence of wall attenuation is rather limited. In this typical office environment, there is also furniture: e.g. metal bookshelves that are about 2 m tall. Please note the selected anchors are not the same as those in figure 2.1. This comes from the fact that the active nodes are also different.

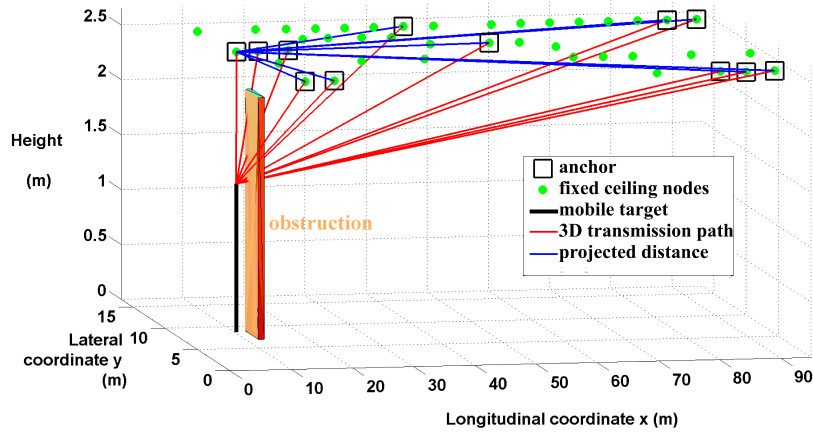


Figure 5.2: Schematic overview of the pseudo-3D environment. All but one pseudo-3D distance (represented by solid red lines) are approximated by their projection in the plane of the fixed nodes (represented by solid blue lines).

Figure 5.2 presents the schematic overview of the pseudo-3D environment. Again, the fixed ceiling nodes are shown as green circles. The remainder of this chapter compares a two-dimensional algorithm with a pseudo-3D algorithm. To this end, two tests are performed. The blue lines belong to the first test in the plane

of the anchors. The red lines belong to the second test, 1.2 m beneath each fixed active node.

- 1 The first test uses 2DLiReFLoA and the two-dimensional plane of the fixed nodes. Please recall from section 4.6.1 2DLiReFLoA is the subsequent implementation of the anchor selection of section 4.1, the preprocessing of section 4.2 and LiReFLoA 4.3. The test conditions are the same as those described at the beginning of section 4.1. For this two-dimensional test, the Euclidean distances are taken into account. In figure 5.2 these distances are represented by blue solid lines.
- 2 The second test is pseudo-3D and uses P3DLiReFLoA. Now, a test person activates a mobile node at a constant height from the ceiling. This test was performed exactly 1.2 m beneath each fixed node. The mobile node is hand carried with the same type inverted-F antenna parallel to the antenna of the fixed node above it, in order to obtain good reproducibility. The swapping time of the sending nodes is different (the target is not fixed on the ceiling, but moved by the person to be located), but the other conditions remain the same. All but one pseudo-3D distance (represented by solid red lines in figure 5.2) are approximated by their projection in the plane of the fixed nodes (represented by solid blue lines). Only for the closest distance, the exact distance is used (1.2 m). Although this introduces an error (of maximum 3.5%) in the calibration of the distances, this error is acceptable because the nodes are on a relatively coarse grid of approximately 4.5 m. This error is an order of magnitude smaller than the error on the multipath faded RSSI-measurements, which has been discussed in section 2.3.2.

5.3 P3DLiReFLoA

Figure 5.3 gives a flowchart of the positioning algorithm. It contains both the 2DLiReFLoA and the new Pseudo-three-Dimensional Linear Regression based Fast Localization Algorithm (P3DLiReFLoA) steps.

The left-hand side illustrates that 2DLiReFLoA is a sequential combination of the algorithms described in section 4.1, section 4.2 and section 4.3. see section 4.6.1

The pseudo-3D algorithm is based on this two-dimensional algorithm. In figure 5.3 the rightmost path is taken. The differences between the two-dimensional and pseudo-3D algorithm are the transmission of the RSSI-beacons beneath the selected anchors and the calibration procedure. The remainder of this section explains this approach.

A full three-dimensional counterpart of the fast two-dimensional anchor selection procedure requires either the automatic height variation of the nodes or,

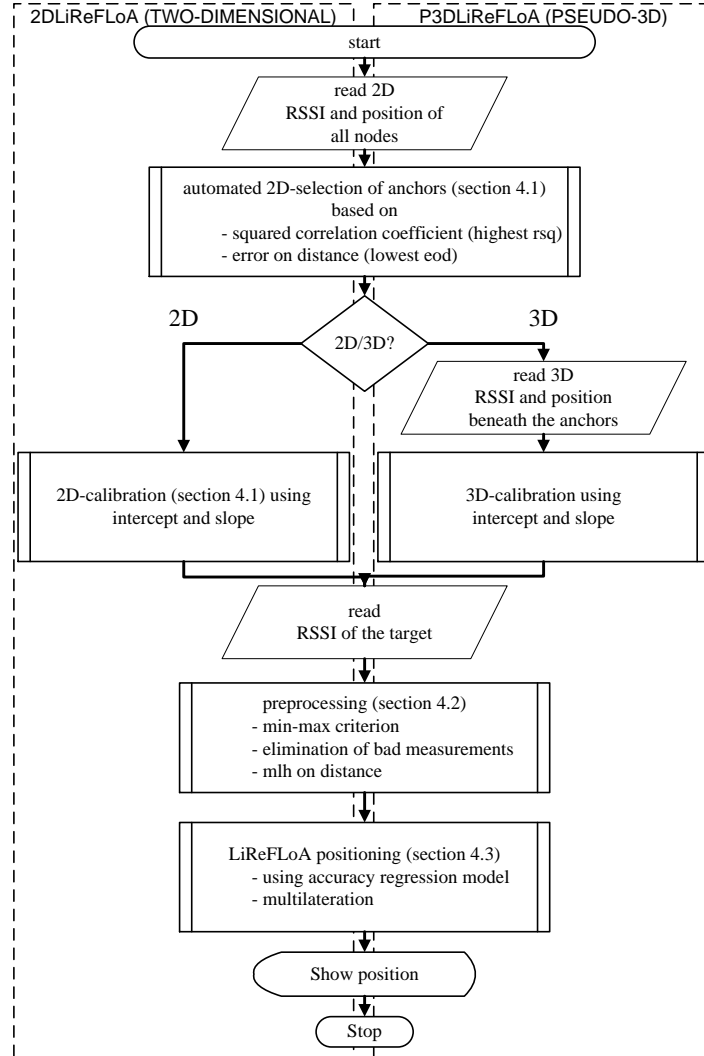


Figure 5.3: This flowchart illustrates the similarities and the differences between the two-dimensional and three-dimensional algorithm.

alternatively, a physical walk in the building with a mobile target transmitting at many known places. Because this full three-dimensional selection is neither cheap nor simple, the two-dimensional selection is kept in our pseudo-3D algorithm.

The two-dimensional calibration of the anchors requires linear regression between the measured RSSI and the logarithm of the distances. The (unknown)

distances can be calculated fast using the intercept and the slope of this regression line. The regression lines change when the target is moved away from the two-dimensional plane of the anchors, basically because there is a larger attenuation due to the presence of furniture. A human body in the vicinity of the transmitter has a comparable attenuating effect on the communication link. Therefore an easy pseudo-3D calibration is performed. It is not necessary to use the complete test data of the second test. Only the test bed RSSI-values of sending mobile node 1.2 m beneath the twelve anchors are needed. A linear regression between the RSSI-values and the logarithmic distance is executed for each of the twelve mobile node positions. The new slopes and intercepts are used for pseudo-3D calibration of the anchors and the distance calculations.

The pseudo-3D selection and calibration has several advantages:

- 1 A good two-dimensional anchor is also a good pseudo-3D anchor, as illustrated in figure 5.4.

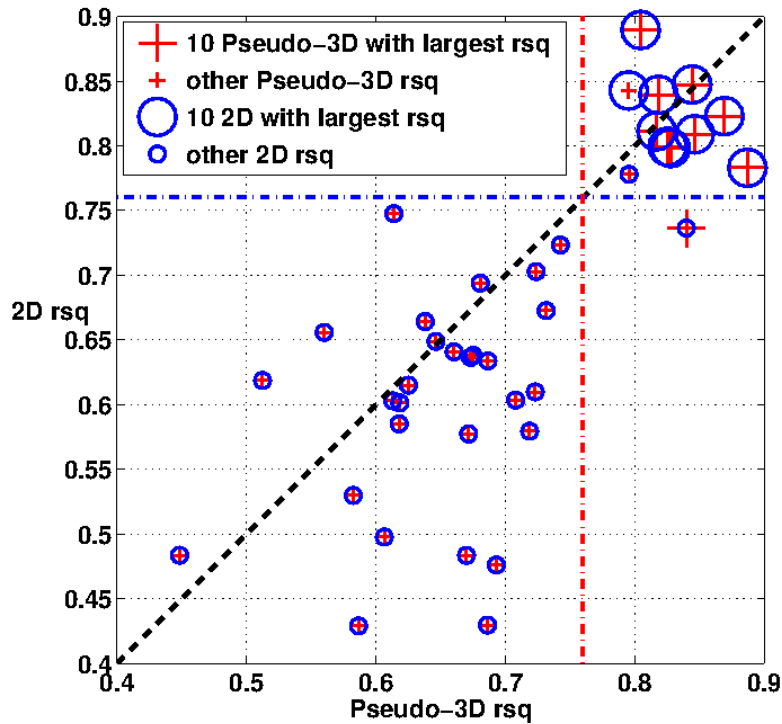


Figure 5.4: Comparison of the correlation coefficient of the two-dimensional and pseudo-3D tests. Two-dimensional nodes with high RSQ are also pseudo-3D nodes with high RSQ.

This figure plots the RSQ-values of the two-dimensional regression (in the first test, i.e. on the ceiling) versus the pseudo-3D regression (in the second test, i.e. with one mobile node for regression). The large circles are the 10 best pseudo-3D correlated counterparts, and the smaller circles represent the others. The large “+”-signs denote the ten best two-dimensional correlated ones and the small “+”-signs are the others. The blue dash-dotted horizontal line corresponds with a two-dimensional RSQ of 0.76 and the red vertical dash-dotted line corresponds with a pseudo-3D RSQ of 0.76. Above this 0.76 value the best correlated nodes are found. The diagonal dashed line represents the bisector of the pseudo-3D RSQ and two-dimensional RSQ axes. The nodes with largest RSQ are in the upper right-hand side part of the figure. A pseudo-3D selection instead of a two-dimensional one would have resulted in only one different anchor at a cost of a higher computational time (because measurements beneath each node are needed). Therefore, in P3DLiReFLoA the two-dimensional selection of the anchors is kept and there is no extra time needed for a dedicated pseudo-3D anchor selection. Please note that a pseudo-3D correlation coefficient of 0.795 (instead of 0.840) is still a very good value. We further observe that the three-dimensional RSQ-values are somewhat higher than the two-dimensional RSQ-values: there are 12 anchors at the right-hand side of the vertical line at 0.76 and only 11 anchors above the horizontal line at 0.76. Furthermore there are 11 points above the bisector (positive ranks), none of the points are on the bisector (no ties) and 30 points below the bisector (negative ranks), therefore a Wilcoxon signed-rank test (see section 4.5.5) [17] rejects the null-hypothesis that the pseudo-3D RSQ-distribution equals the two-dimensional RSQ-distribution.

- 2 Figure 5.3 illustrates that 2DLiReFLoA and P3DLiReFLoA have the same the anchor selection and calibration. Thanks to the automated selection and two-dimensional calibration of nodes, the best available two-dimensional anchors can be selected at a given time. When anchors fail, the algorithm can quickly reselect and calibrate other anchors (see section 4.1.4). This advantage is kept in P3DLiReFLoA. Selecting uniformly distributed anchors could have resulted in an anchor that is down. In our sparse anchor density environment, every selected (and high quality) anchor is needed.
- 3 The measurements beneath an anchor results in a RSSI at a very short distance. This nearby information is very useful, because it is at the beginning of the regression line. Without this measurement, the regression line would have been extrapolated, resulting in large errors [18].
- 4 The complete extension to the pseudo-third dimension requires minimal effort and computation time: only on twelve anchors a linear regression be-

tween the logarithm of the distance and the RSSI is needed to obtain the new propagation constants for the calibration. This can easily be implemented: the person to be located in a building first initiates the pseudo-3D calibration at a limited number of places (beneath the dynamically selected anchors). Afterwards the person can be located anywhere in the building.

The algorithm can calculate the (x,y)-position for any point in the plane 1.2 m beneath the anchors, starting from the position of the anchors and their pseudo-3D propagation parameters. In this chapter the moving target sends packets beneath the deployed two-dimensional nodes, enabling a good comparison with the two-dimensional parameters. Please note that sending at other positions requires an (automated) measurement of the exact position for interpretation of the results of the algorithm. The remainder of the two-dimensional algorithm in figure 5.3 is unchanged in the pseudo-3D algorithm. More about P3DLiReFLoA can be found in [19].

5.4 Results

This section presents the results of the P3DLiReFLoA algorithm. First, the differences between a two-dimensional calibration (with all nodes in the plane of the anchors) and a pseudo-3D calibration (with all nodes 1.2 m beneath the plane of the anchors) is discussed. Next, a comparison is made between P3DLiReFLoA and a maximum likelihood algorithm with a minimum least square error cost function. Finally, a comparison is made between the two-dimensional algorithm (with all targets in the plane of the anchors) and P3DLiReFLoA (with all targets 1.2 m beneath the plane of the anchors).

5.4.1 Comparison of the two-dimensional and pseudo-3D calibration

Figure 5.5 compares the 2DLiReFLoA calibration with the P3DLiReFLoA calibration for a selected anchor (node 56) at the left-hand side of the building (see figure 5.2). The small circles in figure 5.5 are the recorded RSSI-distance pairs for all nodes when the selected anchor is sending (two-dimensional values). The “+”-markers are the recorded RSSI-distance pairs for all nodes when sending with the mobile node beneath that anchor (pseudo-3D values). The solid and the dashed line show the corresponding regression lines. The pseudo-3D regression line has a higher attenuation at low distance levels and a flatter slope. At high distances, the difference in RSSI decreases. This can be explained by the fact that the attenuation of furniture (or other obstacles) is more pronounced at low to medium distances. Please recall the angle dependency of the attenuation factors, discussed in [20]

and figure 4.12: when electromagnetic radiation is obliquely incident on a wall or floor, less power will be transmitted through the wall than would occur at normal incidence. Nodes in the neighborhood of the target have larger incident angles than nodes that are further away. Although only one typical regression comparison is shown here, this conclusion holds true for the vast majority of the nodes, as illustrated in figure 5.6 and figure 5.7.

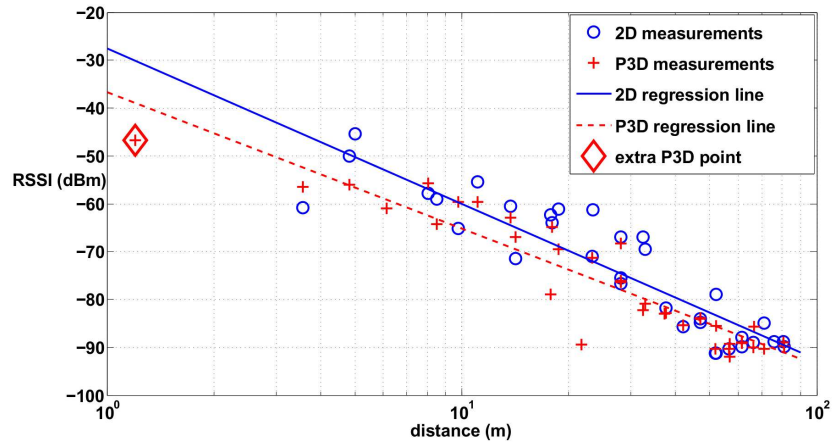


Figure 5.5: Pseudo-3D regression lines not only have higher attenuation at low distances, but are also less steep than two-dimensional counterparts. The pseudo-3D calibration procedure has an extra measurement at the beginning of the regression line.

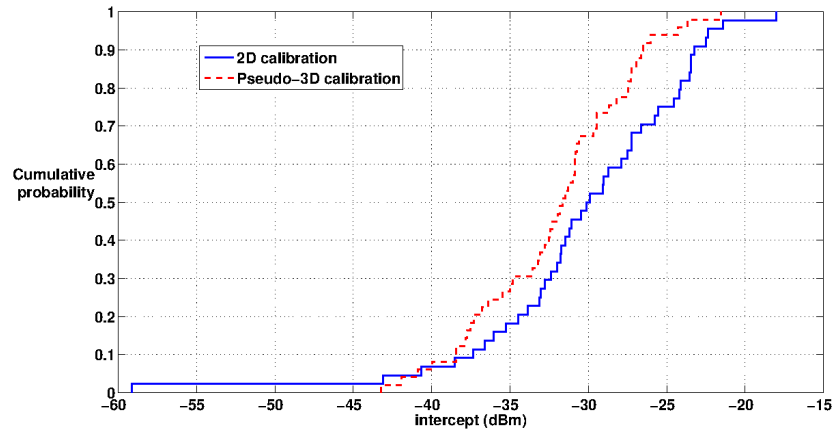


Figure 5.6: The pseudo-3D calibration results in higher attenuation at small distances (compared to two-dimensional calibration).

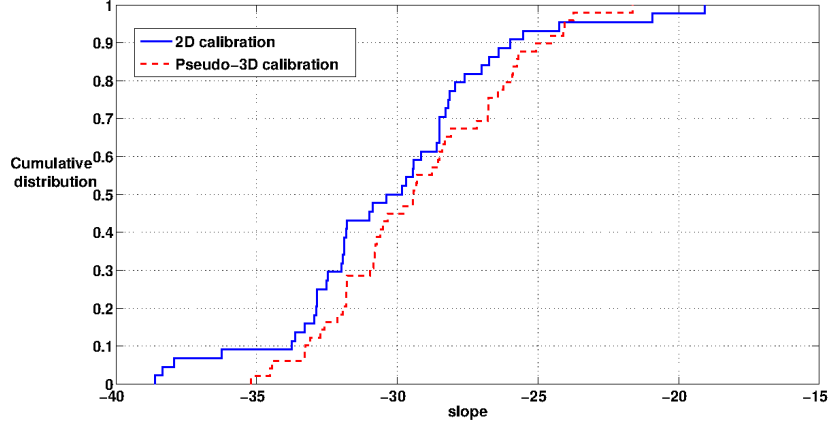


Figure 5.7: The two-dimensional calibration results in higher slopes (compared to pseudo-3D calibration).

Figure 5.6 represents the cumulative distribution plot for the intercept for two-dimensional and pseudo-3D calibrations respectively. Except for a few nodes (not being anchor nodes), the two-dimensional plot is at the right-hand side of the pseudo-3D plot. Hence, this confirms the higher attenuation at low distances of the pseudo-3D calibration compared to the two-dimensional calibration. Figure 5.7 represents the cumulative distribution plot of the slope of the two-dimensional and pseudo-3D calibration of all nodes. Now, the two-dimensional plot is at the left-hand side of the pseudo-3D plot for the vast majority of the nodes. Therefore, the two-dimensional calibration results in higher slopes.

A cumulative distribution function (cdf) plot of the *Error_on_Distance* for the two-dimensional and pseudo-3D calibration of the anchors is presented in figure 5.8. Only the restricted data set limited to the measurements of the anchors (two-dimensional) is considered in this figure. Being defined as twice the estimated standard deviation of the (logarithmic) distance frequency distribution (see section 4.1.2 and [21]), the *Error_on_Distance* is a logarithmic value on the tolerances of the distances. E.g. a value of 0.3 means that the tolerances on a distance are minus $10^{-0.3}$ or minus 50% and plus $10^{+0.3}$ or 200%. This figure illustrates that the anchors have a higher *Error_on_Distance* when pseudo-3D calibrated than when two-dimensionally calibrated. This difference in *Error_on_Distance* can be explained by the higher attenuation in presence of furniture. The tolerances are not only used in the preprocessing step (both elimination of bad measurements and maximum likelihood on the distance), but also in the positioning step for a decision on the amount of constructive multipath fading present. Therefore, it is important to offer the right empirical EOD to the algorithm: the two-dimensional EOD for two-dimensional localization and the pseudo-3D EOD for the pseudo-3D

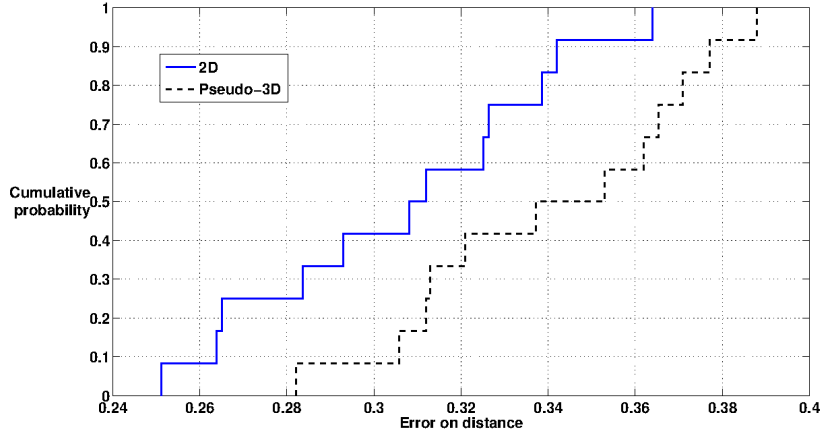


Figure 5.8: The pseudo-3D *Errors_on_Distance* of the anchors are larger than their two-dimensional counterparts.

localization.

5.4.2 Comparison of our algorithm with a more conventional algorithm.

The three-dimensional algorithm is tested: the algorithm is executed and the results are compared with the exact position. For comparison, a maximum likelihood algorithm on the position is implemented using the minimum mean square error (MMSE) cost function (see equation 3.15).

Like in section 3.4, section 4.3.9 and section 4.5, the grid method is used. Please recall that unlike the conjugate gradient algorithm [22], the grid method always finds the exact minimum on the grid and does not get stuck in local minima. Please further recall from section 3.4.2 that the exact minimum on the grid is not always in the immediate vicinity of the wanted minimum. Indeed, figure 3.8 illustrates that in rare cases (and unideal cost functions), the local minimum is closer to the target than the exact minimum on the grid.

The grid plane is now formed 1.2 m below the vertical position of the anchors. The point in this plane with coordinates (x,y) that minimizes the sum of squared differences between the calculated and exact distances is the estimated position.

With both algorithms the position is calculated and compared with the exact distance. This is performed for both cases: with and without our preprocessing steps, which are based on a min-max criterion, elimination of bad measurements and a maximum likelihood on the distance (see section 4.2). The results of the position errors of both algorithms are outlined in table 5.2 and figure 5.9. Table

	P3DLiReFLoA		P3D MMSE	
	w/o (m)	with (m)	w/o (m)	with (m)
Upper outlier	-	-	22.47	17.56
Upper adjacent	19.00	13.37	17.67	12.36
Third quartile	9.80	8.48	10.33	7.68
Median	7.22	5.11	7.79	4.90
First quartile	3.49	2.91	4.70	3.50
Lower adjacent	1.19	0.46	1.79	1.22
Under Outlier	-	-	-	-

Table 5.2: Position error comparison between P3DLiReFLoA and P3D MMSE, both with and without (w/o) the preprocessing.

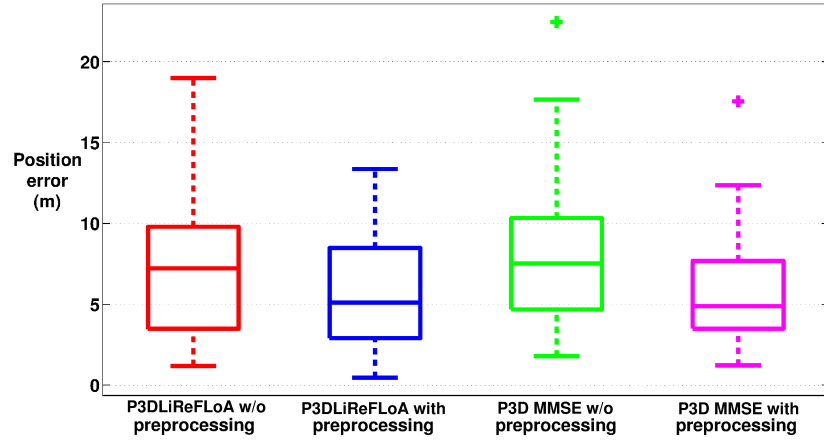


Figure 5.9: Boxplot of the position errors with and without (w/o) the preprocessing steps: at the left-hand side for the position errors of the algorithm proposed in this paper and at the right-hand side for the position errors of the minimum mean square error maximum likelihood algorithm. A maximum likelihood algorithm with a mean square error cost function has a lower position error median when our preprocessing is applied.

5.2 lists the outliers, adjacent (most extreme data points not considered outliers), and the first (25th percentile), second (50th percentile, median) and third (75th percentile) quartile. The same information is given a graphical form with boxplots in figure 5.9. On each box, the central mark is the median, the edges of the box are the first and third quartile, the whiskers extend to the upper and lower adjacent. All position errors that are 1.5 times the box size above the third quartile are deemed upper outliers, and all position errors that are 1.5 times the box size below the first quartile are considered under outliers. The outliers are plotted individually with a

“+”-sign [18]. This table 5.2 and figure 5.9 reveal that

- There are no outliers with P3DLiReFLoA, in contrast with MMSE. For a normal distribution, the edges of the box lie between the average minus 0.6745 times the standard deviation and the average plus 0.6745 times the standard deviation. The box is then 1.349 standard deviations wide. Therefore, the whiskers extend to the average minus/plus 0.6745+1.5x1.349 the standard deviation, or the average minus/plus 2.6980 times the standard deviation. Therefore, the probability that there are values below the lower adjacent is less than $\text{normcdf}(-2.6980)$ or 0.35% in a normal distribution. (Normcdf is the cumulative standard normal distribution function.) Likewise, the probability there are values above the upper adjacent in a normal distribution equals $(1 - \text{normcdf}(2.698))$ or also 0.35%.
- A maximum likelihood algorithm with a mean square error cost function (MMSE) has a higher position error median when our preprocessing is not applied.
- When our preprocessing is applied not only this median, but also the high percentiles of this algorithm are improved.
- None of the distributions are normal distributions:
 - All of them have larger upper tails than lower tails, introducing skewness in the distribution.
 - For MMSE there are outliers, which are absent in normal distributions.
 - Except for MMSE without the processing, the medians (the large horizontal lines in the box) are not in the center of the box. This means that the median of the distribution is different from the average.

Further interpretation of the test results is based on statistical inference. Because the position error is not normally distributed, Students t-tests could lead researchers to draw incorrect conclusions [23]. Non-parametric tests make no assumption on the distribution, and are a better option here. Thanks to the increased availability of software, these non-parametric statistical analyses are often found in medical research [24]. The Wilcoxon signed-rank test (see section 4.5.5 [17]) is a non-parametric statistical hypothesis test for comparing two related samples, e.g. before preprocessing and after. The null-hypothesis of a first Wilcoxon test - that the P3DLiReFLoA position error distribution after the preprocessing equals the P3DLiReFLoA position error distribution before the preprocessing - is rejected. The two-tailed p-value (defined as the probability that the test statistic is equal to or more extreme than the one observed under the null hypothesis) equals 0.2%. A second Wilcoxon test is performed on MMSE with and without the preprocessing

steps and results in a two-tailed p-value of less than 0.1%. The third Wilcoxon test follows, comparing P3DLiReFLoA with the preprocessing steps and MMSE without them, results in a two-tailed p-value of 2%. The first Wilcoxon test concludes that there is a difference in position error for P3DLiReFLoA with and without the preprocessing. Wilcoxon test 2 shows this is also the case for MMSE. This proves our preprocessing has a positive effect on the positioning error. Furthermore, the third Wilcoxon test shows that the position errors of P3DLiReFLoA with the preprocessing steps are significantly lower than those of the MMSE without them.

5.4.3 Comparison of the two-dimensional and pseudo-3D algorithms

This section compares the two-dimensional algorithm (first test, left-hand side of figure 5.3) with the pseudo-3D P3DLiReFLoA algorithm (second test, right-hand side of figure 5.3). Testing the algorithm in their respective two-dimensional and pseudo-3D environment gives comparable medians. Figure 5.10 further illustrates that also the distributions are very similar.

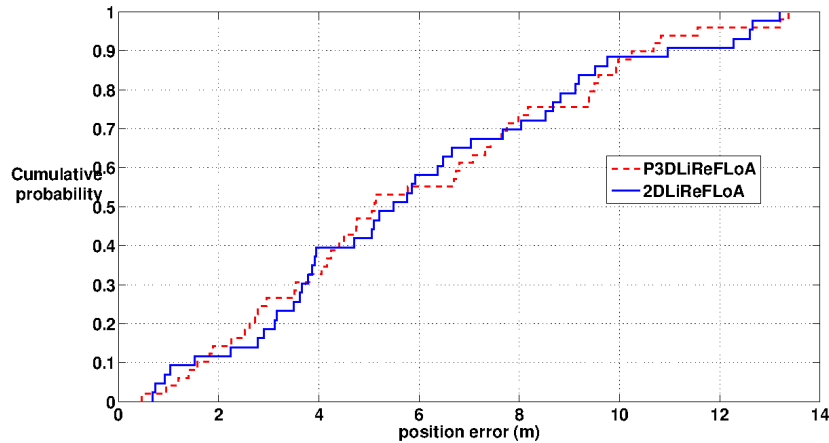


Figure 5.10: Comparison of the cdfplot of the two-dimensional and pseudo-3D algorithms

Table 5.3 gives an overview of the comparison of two-dimensional and pseudo-3D algorithms. This table reveals that both algorithms use the same position engine. Both the preprocessing and the positioning procedure are the same (as already outlined in figure 5.3). Also the anchor selection remains unchanged. While the two-dimensional LiReFLoA based algorithm calibrates these anchors in the two-dimensional plane of these anchors, a pseudo-3D calibration with a mobile node beneath these anchors is needed for P3DLiReFLoA. With an anchor density of 0.008 anchors per square meter, very little additional time is needed. Indeed,

	2DLiReFLoA	P3DLiReFLoA
Localization	Two-dimensional	Pseudo-3D
Anchor selection	Two-dimensional	
Calibration	Two-dimensional	Three-dimensional on selected anchors
Position of target	In anchor plane	1.2 m below anchor plane
Position engine	The same	
Computation time	Fast	Slightly less fast: only quick additional calibration is required for a limited number of anchors
Accuracy		
Without preprocessing	Median 5.79 m	Median 7.22 m
With preprocessing	Median 5.49 m	Median 5.11 m
Cdfplot	See figure 5.10	
Wilcoxon p-values	Two-tailed p-value = 98%	

Table 5.3: Comparison of the two-dimensional and pseudo-3D algorithms

figure 5.3 shows that 2DLiReFLoA and P3DLiReFLoA share the same anchor selection. The person to be located makes an initial walk through the building and beneath the two-dimensionally selected anchors, he triggers the mobile node. A paired Wilcoxon test results in a two-tailed p-value of 98%. Therefore the null-hypothesis that P3DLiReFLoA performs equally well in a pseudo-3D environment as LiReFLoA in a two-dimensional environment is accepted.

5.5 Conclusions

This chapter presents a new pseudo-3D localization algorithm, based on a fast two-dimensional algorithm. Only a quick recalibration is required for the limited number of anchors. Our empirical tests show that the position errors are lower than with a maximum likelihood algorithm with a mean square error cost function. Pre-processing of the data also reduces the position errors for the maximum likelihood algorithm in a statistically significant way.

References

- [1] G. Zanca, F. Zorzi, A. Zanella, and M. Zorzi. *Experimental comparison of RSSI-based Localization Algorithms for Indoor Wireless Sensor Networks*.

- In REALWSN 2008: Proceedings of the workshop on Real-World Wireless Sensor Networks, pages 1–5, Glasgow, Scotland, UK, April 1 2008.
- [2] P. Pivato, L. Palopoli, and D. Petri. *Accuracy of RSS-Based Centroid Localization Algorithms in an Indoor Environment*. IEEE Transactions on Instrumentation and Measurement, 60:3451–3460, 2011.
 - [3] X. Luo, W. J. O'Brien, and C. L. Julien. *Comparative Evaluation of Received Signal-Strength Index (RSSI) based Indoor Localization Techniques for Construction Jobsites*. Journal of Advanced Engineering Informatics, 25:2:355–363, 2010.
 - [4] C. Gao, Z. Yu, Y. Wei, S. Russell S, and Y. Guan. *A Statistical Indoor Localization Method for Supporting Location-based Access Control*. Mobile Networks and Applications, 14:253–263, April 2009.
 - [5] W. S. Cleveland and S. J. Devlin. *Locally-Weighted Fitting: An Approach to Fitting Analysis by Local Fitting*. Journal of the American Statistical Association, 83:403:596–610, 1988.
 - [6] H. Hashemi. *The Indoor Radio Propagation Channel*. In Proceedings of the IEEE Vehicular Technology Conference, pages 416–421, Denver, CO, USA, May 1993.
 - [7] *IEEE Standard for Information Technology – Telecommunications and Information Exchange between Systems – Local and Metropolitan Area Networks – Specific Requirements – Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs)*. <http://standards.ieee.org/getieee802/download/802.15.4-2006.pdf>, 2006. IEEE802.15.4-2006 Std.
 - [8] J. Shirahama and T. Ohtsuki. *RSS-based Localization in Environments with Different Path Loss Exponent for Each Link*. In Proceedings of the IEEE Vehicular Technology Conference, pages 1509–1513, Singapore, May 2008.
 - [9] A. Lewandowski and C. Wietfeld. *A Comprehensive Approach for Optimizing ToA-Localization in Harsh Industrial Environments*. In PLANS2010: IEEE-ION Position Location and Navigation Symposium, pages 516–525, Indian Wells, CA, USA, May 4-6 2010.
 - [10] T. Sarkar, Z. Ji Z, K. Kim KJ, A. Medouri, and M. Salazar-Palma. *A Survey of Various Propagation Models for Mobile Communication*. IEEE Antennas and Propagation Magazine, 45:51–82, 2003.

- [11] D. Plets, W. Joseph, K. Vanhecke, E. Tanghe, L. Martens, S. Bouckaert, I. Moerman, and P. Demeester. *Validation of Path Loss by Heuristic Prediction Tool with Path Loss and RSSI Measurements*. In 2010 IEEE International Symposium on Antennas and Propagation, pages 1–4, Toronto, Canada, July 11–17 2010.
- [12] P. Bahl and V. Padmanabhan. *RADAR: An In-building RF-based User Location and Tracking System*. In Proceedings of the IEEE INFOCOM, volume 2, pages 775–784, Tel-Aviv, Israel, March 2000.
- [13] L. Hamza and C. Nerguizian. *Neural Network and Fingerprinting-based Localization in Dynamic Channels*. In WISP 2009: Proceedings of the 6th IEEE International Symposium on Intelligent Signal Processing, pages 253–258, Budapest, Hungary, August 26–28 2009.
- [14] J. J. Robles, M. Deicke, and R. Lehnert. *3D Fingerprint-based Localization for Wireless Sensor Networks*. In WPNC 2010: Proceeding of the 7th Workshop on Positioning, Navigation and Communication, pages 77–85, Dresden, Germany, March 11–12 2010.
- [15] S. Gansemer, S. Hakobyan, S. Puschel S, and U. Grossmann. *3D WLAN Indoor Positioning in Multi-Storey Buildings*. In IDAACS2009: IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, pages 669–672, Rende, Italy, September 21–23 2009.
- [16] C. Shih and P.J. Marron. *COLA: Complexity-reduced Trilateration Approach for 3D Localization in Wireless Sensor Networks*. In SENSORCOM: Fourth International Conference on Sensor Technologies and Applications, pages 24–32, Venice, Italy, July 18–25 2010.
- [17] J. J. Higgins. *Introduction to Modern nonparametric Statistics*. Duxbury Press, Belmont, CA, USA, 2004.
- [18] J. Neter, W. Wasserman, and M. Kutner. *Applied Linear Statistical Models*. MacGraw Hill, New York, USA, 2005.
- [19] F. Vanheel, J. Verhaevert, E. Laermans, I. Moerman, and P. Demeester. *Pseudo-3D RSSI-based WSN Localization Algorithm using Linear Regression*. Journal on Wireless Communication and Mobile Computing. Article first published online: 23 August 2013. DOI: 10.1002/wcm.2416.
- [20] K. Cheung, J. Sau, and R. Murch. *A New Empirical Model for Indoor Propagation Prediction*. IEEE Transactions on Vehicular Technology, 47:3:996–1001, 1998.

- [21] F. Vanheel, J. Verhaevert, E. Laermans, I. Moerman, and P. Demeester. *Automated Linear Regression Tools Improve RSSI WSN Localization in Multipath Indoor Environment*. EURASIP Journal on Wireless Communications and Networking 2011, Special Issue: Localization in Mobile Wireless and Sensor Networks, 2011:38, July 2011.
- [22] J. L. Nazareth. *Conjugate-gradient Methods*. Wiley Interdisciplinary Reviews: Computational Statistics, 1:348–353, 2009.
- [23] P. Flayers. *Alphas, betas and skewy distributions: two ways of getting the wrong answer*. Advances in Health Sciences Education, 16:3, 2011.
- [24] F. McElduff, M. Cortina-Borja, S. Chan SK, and A. Wade. *When T-Tests or Wilcoxon-Mann-Whitney Tests Won't Do*. Advances in Physiology Education, 34:128–133, 2010.

6

Overall conclusions

6.1 Overall conclusions

With military, law enforcing, firefighting, medical, industrial, civilian and consumer applications, there is a huge potential for localization-aware services. Due to their unique characteristics, receiver signal strength based wireless sensor networks have been called the ideal candidate to fulfill this need. No wonder that this topic received so much research attention during the last few years. Despite the research effort, it is observed that the accuracy of commercial indoor localization systems is still marginal.

Theory and simulations are valuable tools for designing localization algorithms. However, it is important to base these tools on valid assumptions. Therefore, we started this book with the study on the physical layer and its technological challenges. One major revelation is that the underlying physical aspects are sometimes lost in theoretical localization algorithms. Another shows that multipath fading is by far the most troublesome phenomenon for indoor localization, because it generates outliers. A test on simple localization algorithms further showed that there is a trade-off between complexity and robustness to outliers. The Maximum Likelihood algorithm has a high accuracy but is very sensible to outliers. For the Min-Max localization algorithm the opposite is true: a good resilience to outliers is combined with a low accuracy. Therefore, simple localization algorithms need a good preprocessing technique to eliminate outliers.

Until recently, localization experiments were performed on a low scale: a few

nodes, one room, and a few measurements. The wide acceptance of wireless sensor networks and the emergence of test beds added another dimension to this “craft” work: mass-deployment. A main requirement for WSN mass-deployment is the easy installation and configuration. Fingerprinting based solutions today rely on a manual and time-consuming set-up phase. Such solutions cannot deal with fading and the high dynamics of wireless environments (mobility of people, objects and furniture, changing density of people), requiring a recalibration each time the wireless setting changes. In our work, we focus on quasi real-time approaches that can handle the dynamics of the wireless indoor environment. The new developed algorithms in this study are preceded by an automatic selection and calibration of anchor nodes, avoiding complex and time-consuming manual configuration and calibration, while keeping the scalability advantages in number of nodes, space, and number of measurements.

Due to this better scalability, the performance of the statistical tools increases, because more measurements and nodes are encountered. A first reason is that averaging the RSSI-measurements levels out the short-time fading effects. Furthermore, outliers can be identified more accurately in a larger dataset than among just a few measurements.

Algorithms that are more traditional, have a top-down approach: the algorithm is mainly designed on an application level, and what happens on the physical layer is not a major concern. This work presents a bottom-up approach, aimed at keeping the physical aspects into the design. Therefore, the presented preprocessing is based on the underlying physics of the propagation channel, the linear regression model of the RSSI and the logarithmic distance, and on the accuracy of this model.

The common theme through this book is the quest of simple localization algorithms. We presented a regression based anchor selection, calibration and preprocessing. Furthermore, LiReFLoA, LiReCoFuL, 2DLiReFLoA and P3DLiReFLoA have been introduced. Let us further embark on this path.

The conversion of the (corrected) distances (between the anchors and the target) to a position is kept simple in both presented two-dimensional localization algorithms. Linear Regression based Fast Localization Algorithm (LiReFLoA) uses simple geometric principles and further exploits the accuracy of the regression model to reduce the effect of (constructive) multipath fading. Empirical tests in this work show that the combination of the preprocessing and LiReFLoA performs better than the MMSE algorithm. Linear Regression based Cost Function for Localization (LiReCoFuL) is, just like MMSE, a cost function in a maximum likelihood algorithm. The former cost function, however, includes the physical aspects of the power decay with distance and uses linear regression concepts to derive a model that includes the estimation of the standard deviation. Therefore, the position errors are lower than those of the MMSE cost function. Moreover, LiReCoFuL has very few local extremes. This enables it to be used with fewer

convergence problems in a conjugate gradient algorithm.

This thesis ends with the expansion of the two-dimensional LiReFLoA to a pseudo-three-dimensional localization algorithm. These algorithms use two-dimensional projection techniques to find an object in a three-dimensional space. We show with empirical results that P3LiReFLoA is a simple algorithm having lower position errors than those of a maximum likelihood algorithm with a MMSE cost function. Preprocessing the data reduces the position errors of the MMSE algorithm in a statistically significant way.

6.2 Outlook and future work

We feel this thesis has contributed in the design of localization algorithms and should not be considered as an endpoint, but as a source of inspiration for future research work.

Here follows a list of issues, still remaining to be solved and requiring further parallel studies and research.

- The algorithms have been validated on the rectangular second and third floor of the iMinds building. A next logical step is testing the algorithms on non-rectangular shapes. We rejected the idea to virtually increase the width of the building, because this is too artificial. Using a restricted area of the building was also not successful, because this requires a more dense node deployment, in order to keep a minimum quantity of high quality anchors. On square ground floors, we are convinced the anchor selection will give the same conclusions: put anchors at the extremities of the building and add central “low *EOD*” anchors. We think LiReCoFuL will scale perfectly, for LiReFLoA a multilateration in the second dimension might be required.
- In order to meet the different customer requirements, architects can be very creative: e.g. the floor plans of the main hospitals in Ghent, Ypres and Courtray are completely different. Irregularly shaped floor plans can be dealt with by splitting the floor plan into rectangular sections and performing the localization algorithms in each section. Unwanted interaction between the sections can be avoided by a proper channel selection plan.
- Our environment has many drywall walls, a highly conductive ceiling and long corridors. Environments with more concrete walls result in higher RF attenuation, less constructive and more destructive multipath fading. This will probably require a denser anchor deployment.
- The iMinds test bed is an already deployed network. Some sensor networks might want to start from scratch and build the network up in order

to limit the number of nodes. The IEEE 802.15.4 is ideally suited to build ad hoc networks. A solution might be developing a new interactive application. Again, the first anchors are put in the corners. Based on the network connectivity of the corner's anchors this application suggests preliminary positions of additional (central anchors), next the user proceeds to these areas with a potential anchor and gets feedback on a map where the already available anchors locate him. On approval the candidate anchor promotes to a new anchor of the building and the selection is completed with these central nodes. For good algorithm performance estimates of the *slope*, *intercept* and *Error_on_Distance* are needed. These parameters are difficult to predict, because they are strongly environment dependent. Maybe, they can be deducted from the kind of environment: a checklist with the amount of metal, concrete, corridors, ... could result in the recommended values. When the accuracy is insufficient, a temporary network with more nodes needs to be deployed in order to characterize the environment with linear regression. After the anchor selection, calibration and calculation of *Error_on_Distance*, the non-anchor nodes are no longer needed.

- Section 2.3.3 was dedicated to interference. It was shown that different radio emissions are usually unwanted. Radio Interferometric Positioning Systems (RIPS), however, try to benefit from interference. They exploit interfering radio waves from two locations at slightly different frequencies to obtain ranging information for localization. The composite signal has a low beat frequency. The envelope of this signal is used to obtain the position [1–3]. Research on this brand new technique is part of the EVARILOS-project [4].
- Repeatability [5] studies the variation of the measurements: i.e. if the same target is located again at another time, what is the difference on the position. In our work, we retested 2DLiReFLoA on two different days. The first time, we obtained a median position error of 5.29 m. The second time, a median of 5.49 m has been obtained (see table 5.3. In [6] an in-depth study on repeatability is found. We keep this as future work.
- Section 2.3.4 showed that the radiation pattern of an inverted-F has a few dips. Section 3.3.1.2 further added that these dips can cause outliers in simple localization algorithms. Although outliers are dealt with in LiReFLoA, these might be avoided, if the antenna optimization guidelines in [7] are used.
- Section 3.3.2.1 listed an overview of the techniques to solve the set of non-linear localization equations. Applying one of these techniques, e.g. the conjugate gradient method on LiReCoFuL (see section 4.5) can further speed up this algorithm.

- Up to now, our work concentrated on the localization of a target, only using anchors. Furthermore, RSSI-measurements below the noise floor were neglected. The latter implies that the information that the anchor-target link is on a highly attenuated path, is discarded. A next logical step is cooperative localization [8]. In such a setting, localization can be obtained through the cooperation of multiple nodes, using not only the measurements from anchors but also the measurements among pairs of ordinary nodes. Thus, high anchor density or long-range anchor transmissions are no longer required. The additional information gained from these measurements between pairs of ordinary nodes can offer increased accuracy and coverage [9]. This topic can easily be expanded to an algorithm with measurements between all nodes and studied with multivariate statistics.
- Up to now, RSSI-measurements were performed at one single frequency. A next step is using more frequencies to obtain a more accurate RSSI, and hence a more accurate position. This “frequency diversity” [10, 11] has been intensively used to increase the throughput in faded environments, like in [12]. Recently, it can also be found in WSN localization algorithms in these harsh environments, like in [13, 14].
- Environments with large RSSI variability over time need recursive filtering, such as a Kalman filter. A linear Kalman filter is very robust, and with a proper selection of the process noise covariance and the measurement noise covariance convergence is fast, if a node receives data directly from anchors [15]. Furthermore, if the noise is Gaussian and zeromean, the Kalman filter is the optimal estimator in the sense that it minimizes the expected error variance [15, 16]. After the preprocessing steps of section 4.2, the best anchors are selected and the outliers have been deleted. Furthermore, the resulting linear propagation model of the anchors fulfills the Gaussian-ness condition. It looks interesting to insert a Kalman filter after the preprocessing steps in order to continuously follow the changing *slope* and *intercept* of the regression lines.

References

- [1] M. Maróti, P. Völgyesi, S. Dóra, B. Kusý, A. Nádas, A. Lédeczi, G. Balogh, and K. Molnár. *Radio Interferometric Geolocation*. In Proceedings of the 3rd international conference on Embedded networked sensor systems, SENSYS2005, pages 1–12, San Diego, CA, USA, November 2–4 2005.
- [2] B. Kusy, G. Balogh, J. Sallai, A. Lédeczi, and M. Maróti. *InTrack: high precision tracking of mobile sensor nodes*. In Proceedings of the 4th European

- Conference on Wireless Sensor Networks, EWSN2007, pages 51–66, Delft, The Netherlands, January 29–31 2007.
- [3] A. Ledeczi, P. Volgyesi, J. Sallai, and R. Thibodeaux. *A Novel RF Ranging Method*. In International Workshop on Intelligent Solutions in Embedded Systems, pages 1–12, Regensburg, Germany, July 10–11 2008.
 - [4] *Evaluation of RF-based Indoor Localization Solutions for the Future Internet*. http://cordis.europa.eu/search/index.cfm?fuseaction=proj.document&PJ_RCN=13278862.
 - [5] *Standard Practice for use of the Terms Precision and Bias in ASTM Test Methods*. <http://www.astm.org/search/site-search.html?query=E177-13>, 2013. ASTM E177-13 Std.
 - [6] O. Rensfelt, F. Hermans, P. Gunningberg, and L. Larzon. *Repeatable Experiments with Mobile Nodes in a Relocatable WSN Testbed*. In DCOSSW 2010: 6th IEEE International Conference on Distributed Computing in Sensor Systems Workshops, pages 1–6, Santa Barbara, CA, USA, June 21–23 2010.
 - [7] Z. Zhang. *Antenna Design for Mobile Devices*. John Wiley & Sons (Asia), Singapore, 2011.
 - [8] S. Van de Velde, H. Wymeersch, P. Meissner, W. Klaus, and H. Steendam. *Cooperative Multipath-Aided Indoor Localization*. In WCNC 2012: Proceedings of the IEEE Wireless Communications and Networking Conference, pages 3107–3111, Paris, France, April 2012.
 - [9] T. Eren. *Cooperative localization in wireless ad hoc and sensor networks using hybrid distance and bearing (angle of arrival) measurements*. EURASIP Journal on Wireless Communications and Networking 2011, 2011:72, 2011.
 - [10] A. Goldsmith. *Wireless Communications*. Cambridge University Press, 2005.
 - [11] J.G. Proakis. *Digital Communications, Fourth Edition*. MacGraw Hill, New York, USA, 2001.
 - [12] A. Bhartia, Y. Chen, S. Rallapalli, and L. Qiu. *Harnessing Frequency Diversity in Wi-Fi Networks*. In Proceedings of the 17th Annual International Conference on Mobile Computing and Networking, pages 253–264, Las Vegas, Nevada, USA, September 19–23 2011.
 - [13] F. Reichenbach and D. Timmermann. *Indoor Localization with Low Complexity in Wireless Sensor Networks*. In IEEE International Conference on Industrial Informatics, pages 1018–1023, Singapore, August 16–18 2006.

-
- [14] A. Lazaro, D. Girbau, P. Moravek, and R. Villarino. *A Study on Localization in Wireless Sensor Networks using Frequency Diversity for Mitigating Multipath Effects*. Electronics and Electrical Engineering, 19:3, 2013.
 - [15] M. Nilsson. *Localization using Directional Antennas and Recursive Estimation*. In WPNC: Fifth Workshop on Positioning, Navigation and Communication, Hannover, Germany, March 2008.
 - [16] M. S. Grewal and A. P. Andrews. *Kalman Filtering: Theory and Practice Using MATLAB*. John Wiley & Sons, Hoboken, NJ, USA, 2001.

