DISSERTATION:
ON PROXIMITY BASED SUB-AREA LOCALIZATION

by

Aylin Korkmaz

MS, Sabanci University, 2006

Submitted to the Graduate Faculty of

Telecommunications and Networking Program in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

SCHOOL OF INFORMATION SCIENCES

This dissertation was presented

by

Aylin Korkmaz

It was defended on

July 20th 2011

and approved by

Assoc. Prof. Prashant Krishnamurthy, University of Pittsburgh

Assoc. Prof. David Tipper, University of Pittsburgh

Assoc. Prof. Martin B. H. Weiss, University of Pittsburgh

Assist. Prof. Konstantinos Pelechrinis, University of Pittsburgh

Assoc. Prof. Nirupama Bulusu, Portland State University

Dissertation Director: Assoc. Prof. Prashant Krishnamurthy, University of Pittsburgh

DISSERTATION:
ON PROXIMITY BASED SUB-AREA LOCALIZATION

Aylin Korkmaz, PhD

University of Pittsburgh, 2011

A localization system can save lives in the aftermath of an earthquake; position people or valuable assets during a fire in a building; or track airplanes besides many of its other attractive applications. Global Positioning System (GPS) is the most popular localization system, and it can provide 7-10 meters localization accuracy for outdoor users; however, it has certain drawbacks for indoor environments. Alternatively, wireless networks are becoming pervasive and have been densely deployed for communication of various types of devices indoors, exploiting them for the localization of people or other assets is a convenience. Proximity based localization that estimates locations based on closeness to known reference points, coupled with a widely deployed wireless technology, can reduce the cost and effort for localization in local and indoor areas. In this dissertation, we propose a proximity based localization algorithm that exploits knowledge of the overlapping coverages of known monitoring stations. We call this algorithm Sub-Area Localization (SAL). We present a systematic study of proximity-based localization by defining the factors and parameters that affect the localization performance in terms of metrics such as accuracy and efficiency. Then, we demonstrate that SAL can be used in multi-floor buildings to take advantage of the infrastructure elements deployed across floors to reduce the overall cost (in terms of the number of monitoring stations required) without harming accuracy. Finally, we present a case study of how SAL can be used for spatial spectrum detection in wireless cognitive networks.

TABLE OF CONTENTS

PREFACE xx

1.0 INTRODUCTION 1

1.1 Challenges and Motivations 2

1.2 Problem Statement 3

1.3 Contributions 5

2.0 LITERATURE REVIEW 8

2.1 Localization Systems and Performance Metrics 8

2.2 Localization Algorithms 10

2.2.1 Coarse Grained Localization (Proximity-based) 10

2.2.2 Fine-Grained Localization 13

2.2.2.1 Distance-Based Approach: 13

2.2.2.2 Angle-based approach: 19

2.2.2.3 Static Scene Analysis Approach: 20

2.3 RSS Based Localization 21

2.3.1 Calibration-based Localization 21

2.3.1.1 Why Calibration? 21

2.3.1.2 Localization Technique 23

2.3.1.3 Deterministic Calibration-based Approaches 24

2.3.1.4 Probabilistic Calibration-Based Approaches 28

2.3.2 Calibration-Free Localization 32

2.3.2.1 Deterministic Calibration-free Approaches 33

2.3.2.2 Probabilistic Calibration-Free Approaches 38

2.3.3 Comparison of Calibration-based and Calibration-free Localization . 42
2.4 Reference Node Placement Strategies 49
2.5 Multiple Floor Localization Algorithms 50
3.0 SUB-AREA LOCALIZATION ALGORITHM 53
3.1 Sub-Area Localization (SAL) 56
3.1.1 Accuracy 57
3.1.2 Availability 59
3.1.3 Efficiency 61
3.2 Centroid Algorithm (CA) 61
3.3 Performance Under Ideal Channel Conditions 63
3.3.1 Subarea Characterization 64
3.3.2 Heuristics Towards the Best Monitoring Range 67
3.3.3 Evaluation of Localization Error 75
3.3.4 Evaluation of Efficiency 78
3.4 Impact of Log-Normal Shadowing 80
3.4.1 Effect of Wall Attenuation Factor 87
3.4.2 Sensitivity to Multiple Detection Thresholds 91
3.4.3 Incremental Addition of Monitoring Stations 94
3.5 Perturbation from Grid Placement 96
3.6 Performance under Random Placement 102
3.7 Conclusions 107
4.0 MULTI-FLOOR SUB-AREA LOCALIZATION 114
4.1 Multi-floor Sub-area Localization (MSAL) 117
4.1.1 Adjacent Monitoring Range 117
4.1.1.1 Using Isotropic Antenna Pattern 119
4.1.1.2 Using Modified Dipole Antenna Pattern 119
4.1.2 Utilizing Multi-Floor MoS Sets 121
4.1.2.1 2FRS ($< (5,4), \mathbf{G}_p, A >$) 122
4.1.2.2 3FRS ($< (4,3,2), \mathbf{G}_p, A >$) 124
4.1.3 Availability 124

4.1.4 Efficiency . . . 126
4.2 Multi-Floor Centroid Algorithm (MCA) . . . 127
4.3 Analysis and Evaluation of MSAL . . . 127
4.3.1 Performance under Ideal Channel Conditions . . . 127
4.3.1.1 2FRS ($< (5,4), \mathbf{G}_p, A >$) . . . 128
4.3.1.2 3FRS ($< (4,3,2), \mathbf{G}_p, A >$) . . . 134
4.3.2 Impact of Log-Normal Shadowing . . . 139
4.3.3 Effect of Wall Attenuation Factor . . . 149
4.3.4 Effect of Floor Attenuation Factor . . . 159
4.3.5 Effect of Random Placement . . . 169
4.4 Floor Determination for MSAL . . . 179
4.4.1 Majority Based Floor Determination (MBFD) . . . 179
4.4.1.1 Special case . . . 180
4.4.2 False Negatives . . . 181
4.4.3 Effect of Antenna Type . . . 181
4.4.4 Evaluation of MBFD Algorithm . . . 182
4.5 Evaluation of MSAL with MBFD . . . 189
4.5.1 Availability with MBFD . . . 190
4.5.2 Best Accuracy Conditions . . . 194
4.5.3 Best Efficiency Conditions vs. Floor Height . . . 197
4.5.4 Best Efficiency Conditions vs. Shadowing Effect . . . 198
4.5.5 Effect of Floor Attenuation Factor . . . 201
4.5.5.1 Best Accuracy Conditions . . . 201
4.5.5.2 Best Efficiency Conditions vs. Floor Height . . . 204
4.5.5.3 Best Efficiency Conditions vs. Shadowing Effect . . . 206
4.6 Conclusion . . . 211
5.0 CASE STUDY: USING SAL FOR SPATIO-TEMPORAL SENSING . . . 215
5.1 Related Work . . . 217
5.2 The Setting for Secondary Spectrum Use with SAL . . . 218
5.2.1 Global Spatio-temporal Sensing with SAL . . . 219

5.3 Examination of the Potential of SAL for Secondary Spectrum Use 221
5.3.1 Evaluation of Efficiency of SAL in Detecting Primary Transmitters . 221
5.3.2 Effect of Frequency Offset on Location Estimation 228
5.3.2.1 Free Space Path Loss Model 228
5.3.2.2 Near Building Height Path Loss Model 233
5.3.3 Effect of Frequency Offset on Spectrum Hole Detection 236
5.4 Discussion on Possible Research Directions 238
6.0 CONCLUSIONS AND FUTURE WORK 241
6.1 Future Work 243
APPENDIX. DETAILED SIMULATION RESULTS FOR VARIOUS LOCALIZATION SCENARIOS 246
A.1 MSAL 246
A.1.1 Effect of Wall Attenuation Factor 246
A.1.2 Effect of Random Placement 246
A.1.3 Summary of Results 252
A.2 Multi-Floor Localization & Floor Determination with 2FRS 273
A.3 Multi-Floor Localization & Floor Determination with 3FRS 275
A.3.1 Availability with MBFD 286
A.3.2 Best Accuracy Conditions 290
A.3.3 Best Efficiency Conditions vs. Floor Height 296
A.3.4 Best Efficiency Conditions vs. Shadowing Effect 304
A.3.5 Effect of Floor Attenuation Factor 307
A.3.5.1 Best Accuracy Conditions 307
A.3.5.2 Best Efficiency Conditions vs. Floor Height 310
A.3.5.3 Best Efficiency Conditions vs. Shadowing Effect 312
A.4 Case Study 318
BIBLIOGRAPHY 321

LIST OF TABLES

3.1 SAL acronyms and variables 62

3.2 Number of distinct subareas A_i as a function of d; $N = 4$ MoSs 67

3.3 Number of distinct subareas A_i as a function of d; $N = 9$ MoSs 68

3.4 Heuristics for $\hat{d}_i$, $i = 1, 2, 3$ for $N = 4$ and $N = 9$ 74

3.5 Optimal monitoring range $\hat{d}$ and minimum localization error 76

3.6 SAL accuracy with $N = 4$ MoSs in various localization scenarios 110

3.7 SAL+CA accuracy with $N = 4$ MoSs in various localization scenarios 111

3.8 CA accuracy with $N = 4$ MoSs in various localization scenarios 112

4.1 Minimum localization error in 2FRS, $\sigma = 0$ 129

4.2 Minimum localization error conditions in 3FRS 138

4.3 MSAL acronyms and variables 140

4.4 Availability at minimum error conditions ($h = 0.25L$) 143

5.1 MoS ranges used in calculating enlarged regions with 4 MoSs 225

A1 MSAL accuracy with 2FRS in various localization scenarios 253

A2 MSAL efficiency with 2FRS in various localization scenarios 254

A3 MSAL accuracy with 3FRS in various localization scenarios 255

A4 MSAL efficiency with 3FRS in various localization scenarios 256

A5 MSAL+MCA accuracy with 2FRS in various localization scenarios 257

A6 MSAL+MCA efficiency with 2FRS in various localization scenarios 258

A7 MSAL+MCA accuracy with 3FRS in various localization scenarios 259

A8 MSAL+MCA efficiency with 3FRS in various localization scenarios 260

A9 MCA-iso accuracy with 2FRS in various localization scenarios 261

A10 MCA-iso efficiency with 2FRS in various localization scenarios 262
A11 MCA-iso accuracy with 3FRS in various localization scenarios 263
A12 MCA-iso efficiency with 3FRS in various localization scenarios 264
A13 MSAL+MCA-dpl accuracy with 2FRS in various localization scenarios 265
A14 MSAL+MCA-dpl efficiency with 2FRS in various localization scenarios . . . 266
A15 MSAL+MCA-dpl accuracy with 3FRS in various localization scenarios 267
A16 MSAL+MCA-dpl efficiency with 3FRS in various localization scenarios . . . 268
A17 MCA-dpl accuracy with 2FRS in various localization scenarios 269
A18 MCA-dpl efficiency with 2FRS in various localization scenarios 270
A19 MCA-dpl accuracy with 3FRS in various localization scenarios 271
A20 MCA-dpl efficiency with 3FRS in various localization scenarios 272
A21 Minimum localization error in 3FRS, $< (5,3,3), \mathbf{G}_p, A >$, $h = 0.25L$ $\sigma = 0$ . . 296

LIST OF FIGURES

2.1 Proximity-based localization 12
2.2 Lateration (a) with no errors, (b) with errors 15
2.3 Triangulation 20
2.4 Histograms of the RSS from an AP, observed at a mobile station (a) closer to the AP and (b) away from the AP 22
2.5 RSS distance curves for AP_2 when there are four APs in total 34
2.6 Triangular interpolation [1] 35
2.7 Triangular extrapolation [1] 35
2.8 A simple acyclic directed graph model [2] 40
2.9 Nonhierarchical Bayesian Graphical Model [2] 42
2.10 Hierarchical Bayesian Graphical Model [2] 43
2.11 Median error per experiment area for calibration-based and calibration-free techniques 44
2.12 Error metric vs. the number of training locations in RADAR [3] 45
2.13 Logarithm of error metric for calibration-based techniques 46
2.14 Logarithm of error metric for calibration-based and calibration-free techniques, for changing values of j 47
2.15 Logarithm of error metric for calibration-based and calibration-free techniques, for changing values of k 47
2.16 Calibration-based and Calibration-free Localization Systems 52
3.1 The sub-area localization 57
3.2 Real and unit disk ranges 59

3.3 Visualization of subareas in localization scenario $< 4, \mathbf{G}, A >$ 65
3.4 Visualization of subareas in localization scenario $< 9, \mathbf{G}, A >$ 66
3.5 Visualization of regions in A when $\sqrt{2}L/2 < d < L$ 69
3.6 Ratio of K to the maximum area among all R_i's for various d/L 69
3.7 Visualization of subareas when $N = 9$, $\sqrt{2}L/4 < d < L/2$ 70
3.8 Areas of regions R_i in % of A as a function of d/L, $N = 4$ 71
3.9 Areas of regions R_i in % of A as a function of d/L, $N = 9$ 72
3.10 Visualization of diagonals and intersection points in A when $\sqrt{2}L/2 < d < L$ 72
3.11 Average and maximum error in $< 4, \mathbf{G}, A >$ and $< 9, \mathbf{G}, A >$ 75
3.12 CDF of localization error 77
3.13 Subareas for smallest localization error with $N = 25$ MoSs 77
3.14 The smallest mean and maximum localization errors vs. N and curve fits . . 78
3.15 Best values of d (that minimizes the maximum error) vs. N 78
3.16 Efficiency $\eta(i, j)$ of SAL 79
3.17 Minimum average error variance with $< 4, \mathbf{G}, A, sh >$ and $< 9, \mathbf{G}, A, sh >$. 83
3.18 Minimum E_{avg} with $< 4, \mathbf{G}, A, sh >$ and $< 9, \mathbf{G}, A, sh >$ 83
3.19 $\frac{d^*}{L} = argmin(E_{avg})$ with $< 4, \mathbf{G}, A, sh >$ and $< 9, \mathbf{G}, A, sh >$ 84
3.20 Availability at $d^* = argmin(E_{avg})$ with $< 4, \mathbf{G}, A, sh >$ and $< 9, \mathbf{G}, A, sh >$. 84
3.21 η at $d^* = argmin(E_{avg})$ with $< 4, \mathbf{G}, A, sh >$ and $< 9, \mathbf{G}, A, sh >$ 85
3.22 Maximum availability with $< 4, \mathbf{G}, A, sh >$ and $< 9, \mathbf{G}, A, sh >$ 85
3.23 $\frac{d^*}{L} = argmax(R)$ with $< 4, \mathbf{G}, A, sh >$ and $< 9, \mathbf{G}, A, sh >$ 86
3.24 Average error at $\frac{d^*}{L} = argmax(R)$ with $< 4, \mathbf{G}, A, sh >$ and $< 9, \mathbf{G}, A, sh >$. 86
3.25 Average error at $\frac{d^*}{L} = argmax(R)$ with $< 4, \mathbf{G}, A, sh >$ and $< 9, \mathbf{G}, A, sh >$. 87
3.26 (a) Floor Plan 1 (FP 1) (b) Floor Plan 2 (FP 2) 89
3.27 Minimum average error with $< 4, \mathbf{G}, A, waf >$ on FP1 and FP2 89
3.28 $\frac{d^*}{L} = argmin(E_{avg})$ with $< 4, \mathbf{G}, A, waf >$ on FP1 and FP2 90
3.29 Availability at d^* with $< 4, \mathbf{G}, A, waf >$ on FP1 and FP2 90
3.30 Efficiency at d^* with $< 4, \mathbf{G}, A, waf >$ on FP1 and FP2 91
3.31 Visualization of subareas using three RSS thresholds with $< 4, \mathbf{G}, A >$ 92
3.32 Minimum E_{avg} using multiple RSS thresholds with $< 4, \mathbf{G}, A, sh >$ 92

3.33 $\frac{d^*}{L} = argmin(E_{avg})$ using multiple RSS thresholds with $< 4, \mathbf{G}, A, sh >$ . . . 93
3.34 Availability at $\frac{d^*}{L}$ using multiple RSS_{th}s with $< 4, \mathbf{G}, A, sh >$ . . . 93
3.35 Efficiency at $\frac{d^*}{L}$ using multiple RSS_{th}s with $< 4, \mathbf{G}, A, sh >$ . . . 94
3.36 Minimum average error with $< 4, \mathbf{G}, A, sh >$ . . . 95
3.37 $\frac{d^*}{L} = argmin(E_{avg})$ with $< 4, \mathbf{G}, A, sh >$ . . . 95
3.38 Availability at $d^* = argmin(E_{avg})$ with $< 4, \mathbf{G}, A, sh >$ . . . 96
3.39 η at $d^* = argmin(E_{avg})$ with $< 4, \mathbf{G}, A, sh >$ . . . 97
3.40 Perturbation on grid, $N = 4$ . . . 98
3.41 Minimum average error with $< 4, \tilde{G}, A, sh >$ . . . 99
3.42 Availability at $d^* = argmin(E_{avg})$ with $< 4, \tilde{G}, A, sh >$ . . . 99
3.43 Efficiency at $d^* = argmin(E_{avg})$ with $< 4, \tilde{G}, A, sh >$ . . . 100
3.44 Maximum availability with $< 4, \tilde{G}, A, sh >$ . . . 101
3.45 $\frac{d^*}{L} = argmax(R)$ with $< 4, \tilde{G}, A, sh >$ . . . 101
3.46 Average error at $\frac{d^*}{L} = argmax(R)$ with $< 4, \tilde{G}, A, sh >$ . . . 102
3.47 Efficiency at $\frac{d^*}{L} = argmax(R)$ with $< 4, \tilde{G}, A, sh >$ . . . 102
3.48 Random placement of $N = 4$ MoSs in $L \times L$ . . . 104
3.49 Minimum E_{avg} with $< 4, \mathbf{R}, A, sh >$ and $< 8, \mathbf{R}, A, sh >$ . . . 105
3.50 $d^* = argmin(E_{avg})$ with $< 4, \mathbf{R}, A, sh >$ and $< 8, \mathbf{R}, A, sh >$ . . . 106
3.51 Availability at $d^* = argmin(E_{avg})$ with $< 4, \mathbf{R}, A, sh >$ . . . 106
3.52 Maximum R with $< 4, \mathbf{R}, A, sh >$ . . . 107
3.53 $d^* = argmax(R)$ with $< 4, \mathbf{R}, A, sh >$ . . . 108
3.54 E_{avg} at $d^* = argmax(R)$ with $< 4, \mathbf{R}, A, sh >$ . . . 109
3.55 How does SAL compare to fine grained localization? . . . 113
4.1 Range of an MoS in its own floor and in adjacent floors, (a) Isotropic antenna, (b) Dipole antenna . . . 118
4.2 Range of an MoS in its own floor and in adjacent floors with Isotropic antenna 120
4.3 Range of an MoS in its own floor and in adjacent floors with Dipole antenna 120
4.4 (a) 3FRS, $< (4, 3, 2), \mathbf{G}_p, A >$, (b) 2FRS, $< (5, 4), \mathbf{G}_p, A >$ . . . 122
4.5 Subareas on floors 1 and 2 with MSAL $< (5, 4), \mathbf{G}_p, A >$ and SAL $< 5, \mathbf{G}, A >$ and SAL $< 4, \mathbf{G}, A >$ . . . 123

4.6 Ranges of MoSs in two floors for barely covering a floor. 124
4.7 Subareas on floors 2,3 and 4 with MSAL $< (4,3,2), \mathbf{G}_p, A >$ and SAL $<$ $3, \mathbf{G}, A >$, SAL $< 2, \mathbf{G}, A >$, and SAL $< 4, \mathbf{G}, A >$ 125
4.8 Localization errors on floors 1 and 2 with MSAL-iso and MSAL-dpl, $\sigma = 0$, $h = 0.25L$ 130
4.9 Sub-areas in floors 1 and 2 for $E^*_{avg} = 0.081L$, $\sigma = 0$ 131
4.10 Effect of monitoring ranges and floor height on availability 132
4.11 Effect of floor height on E^*_{avg} with 2FRS 133
4.12 Best monitoring ranges $d^*_{ii}/L,\ i = 1, 2$ with 2FRS 133
4.13 $\frac{d^*_{11}}{d^*_{22}}$ with 2FRS 134
4.14 E_{avg} in 3FRS with isotropic and dipole antenna, $\sigma = 0$, $h = 0.25L$ 135
4.15 E_{avg} for 3FRS with isotropic antenna, $\sigma = 0$, $h = 0.25L$ 136
4.16 E_{avg} for 3FRS with dipole antenna, $\sigma = 0$, $h = 0.25L$ 137
4.17 Effect of floor height on E^*_{avg} with 3FRS 138
4.18 Best monitoring ranges $d^*_{ii}/L,\ i = 1, 2$ with 3FRS 139
4.19 Effect of shadowing on E^*_{avg} with 2FRS 141
4.20 Effect of shadowing on E^*_{avg} with 3FRS 141
4.21 Availability at d^*_{ii} 2FRS with isotropic and dipole antenna, $h = 0.25L$ 142
4.22 Availability at d^*_{ii}, 3FRS with isotropic and dipole antenna, $h = 0.25L$ 142
4.23 Effect of σ on d^*_{ii} with 2FRS 144
4.24 Effect of σ on d^*_{ii} with 3FRS 145
4.25 E^*_{avg} 2FRS with isotropic and dipole antenna for different values of h and σ . 146
4.26 E^*_{avg} 3FRS with isotropic and dipole antenna for different values of h and σ . 146
4.27 E^*_{avg} 2FRS with isotropic and dipole antenna, $\sigma = 5$ 147
4.28 Efficiency at d^* with 2FRS and 3FRS, $h = 0.25L$ 147
4.29 R^*_{avg} 2FRS with isotropic and dipole antenna 148
4.30 R^*_{avg} 3FRS with isotropic and dipole antenna 148
4.31 Maximum achievable efficiency metric with 2FRS at $\sigma = 0$ and $\sigma = 5$ 150
4.32 Maximum achievable efficiency metric with 3FRS at $\sigma = 0$ and $\sigma = 5$ 150
4.33 Maximum achievable efficiency metric with 2FRS and 3FRS 151

4.34 Accuracy versus h at $\sigma = 0$ with 2FRS 153
4.35 Minimum error versus h at $\sigma = 0$ with 3FRS 154
4.36 Visualization of estimated locations with 2FRS with Floor Plan 1, $h = 0.25L$ 155
4.37 Availability at minimum error conditions versus h at $\sigma = 0$ with 2FRS 156
4.38 Availability at minimum error conditions versus h at $\sigma = 0$ with 3FRS 157
4.39 Minimum error monitoring range versus h at $\sigma = 0$ with 2FRS 158
4.40 Maximum efficiency versus h at $\sigma = 0$ with 2FRS 160
4.41 Maximum efficiency versus h at $\sigma = 0$ with 3FRS 161
4.42 Minimum error for different values of h at $\sigma = 5$ with 2FRS 162
4.43 Minimum error for different values of h at $\sigma = 5$ with 3FRS 163
4.44 Maximum efficiency for different values of σ and $h = 0.25L$ with 2FRS 164
4.45 Maximum efficiency for different values of σ and $h = 0.25L$ with 3FRS 165
4.46 Effective floor height when $FAF = 13$dB 166
4.47 Minimum error for different values of h at $\sigma = 0$ with 2FRS 167
4.48 Minimum error for different values of h at $\sigma = 0$ with 3FRS 167
4.49 Minimum error monitoring range for different values of h at $\sigma = 0$ with 2FRS 168
4.50 Minimum error monitoring range for different values of h at $\sigma = 0$ with 3FRS 169
4.51 Minimum error for different values of h at $\sigma = 5$ with 2FRS 170
4.52 Minimum error for different values of h at $\sigma = 5$ with 3FRS 170
4.53 Availability at minimum error conditions for different values of h at $\sigma = 5$ with 2FRS 171
4.54 Availability at minimum error conditions for different values of h at $\sigma = 5$ with 3FRS 171
4.55 Minimum error monitoring range for different values of h at $\sigma = 5$ with 2FRS 172
4.56 Minimum error monitoring range for different values of h at $\sigma = 5$ with 3FRS 172
4.57 Minimum error monitoring range for different values of s with 2FRS 173
4.58 Minimum error monitoring range for different values of s with 2FRS 174
4.59 Minimum error monitoring range for different values of s with 3FRS 175
4.60 Minimum error monitoring range for different values of s with 3FRS 176

4.61 Accuracy of different instances of random placement with 2FRS, $< (5,4), \mathbf{R}, A >$, $\sigma = 0$ . . . 177
4.62 Accuracy with 2FRS, $< (5,4), \mathbf{R}, A >$ . . . 178
4.63 Efficiency with 2FRS, $< (5,4), \mathbf{R}, A >$ . . . 178
4.64 Second constraint for dipole antennas . . . 184
4.65 Percentage of false negatives with $< (5,4), \mathbf{G}_p, A, iso >$, $\sigma = 0$, $h = 0.5L$ . . . 185
4.66 percentage of false negatives with $< (5,4), \mathbf{G}_p, A, iso >$, $\sigma = 0$, $h = 0.25L$ . . . 186
4.67 Minimum false negative percentage with $< (5,4), \mathbf{G}_p, A, iso >$ . . . 187
4.68 Minimum false negative percentage with $< (5,4), \mathbf{G}_p, A, dpl >$ . . . 188
4.69 Minimum false negative percentage with $< (5,4), \mathbf{G}_p, A, iso/dpl >$, $\sigma = 0$ . . . 189
4.70 Minimum false negative percentage with $< (5,4), \mathbf{G}_p, A, iso/dpl >$, $h = 0.25L$ 190
4.71 2FRS Average availability with different decision depths with isotropic antennas, $\sigma = 0$, $h = 0.25L$ . . . 192
4.72 2FRS Average availability with different decision depths with dipole antennas, $\sigma = 0$, $h = 0.25L$ . . . 193
4.73 Minimum localization error with respect to h when $\sigma = 0$, 2FRS . . . 195
4.74 Minimum localization error with respect to h when $\sigma = 5$, 2FRS . . . 196
4.75 Maximum efficiency with respect to h when $\sigma = 0$, 2FRS . . . 197
4.76 Maximum efficiency with respect to h when $\sigma = 5$, 2FRS . . . 198
4.77 Maximum efficiency with respect to σ when $h = 0.25L$, 2FRS . . . 199
4.78 Average localization error at maximum efficiency conditions with respect to σ when $h = 0.25L$, 2FRS . . . 200
4.79 Availability at maximum efficiency conditions with respect to σ when $h = 0.25L$, 2FRS . . . 201
4.80 Minimum localization error with respect to h when $\sigma = 0$, 2FRS . . . 202
4.81 False negative percentage with respect to h when $\sigma = 0$, 2FRS . . . 203
4.82 Minimum localization error with respect to h when $\sigma = 5$, 2FRS . . . 203
4.83 False negative percentage with respect to h when $\sigma = 5$, 2FRS . . . 204
4.84 Maximum efficiency with respect to h when $\sigma = 0$, 2FRS . . . 205
4.85 Maximum efficiency with respect to h when $\sigma = 5$, 2FRS . . . 206

4.86 Availability with respect to h when $\sigma = 0$, 2FRS 207
4.87 Availability with respect to h when $\sigma = 5$, 2FRS 207
4.88 Maximum efficiency with respect to σ when $h = 0.25L$, 2FRS 208
4.89 Localization error with respect to σ when $h = 0.25L$, 2FRS 209
4.90 False negative percentage with respect to σ when $h = 0.25L$, 2FRS 210
5.1 SAL usage in CWN 220
5.2 Worst case position for primary receiver 220
5.3 Visualization of regions in A when $\sqrt{2}L/2 < d < L$ 223
5.4 Visualization of extended region R_2' 224
5.5 Intersection points of three circles C_2, C_3 and C_4 226
5.6 $N = 4$ MoSs, η_{sh} 227
5.7 $N = 9$ MoSs, η_{sh} 228
5.8 Spectrum hole detection efficiency, η_{sh}^* 229
5.9 Best monitoring range at η_{sh}^* 229
5.10 Performance when optimized for LS, $N = 4$ MoSs, free space path loss model, $f_{LS} = 900$MHz 231
5.11 Performance when optimized for LS, $N = 4$ MoSs, free space path loss model, $f_{LS} = 1900$MHz 232
5.12 Performance when optimized for LS, $N = 4$ MoSs, building height, $f_{LS} = 900$MHz 234
5.13 Performance when optimized for LS, $N = 4$ MoSs, building height, $f_{LS} = 1900$MHz 235
5.14 Change of monitoring range with frequency offset 236
5.15 Visualization of regions when $f_{LS} = 900$MHz and $f = 700$MHz 237
5.16 Visualization of regions when $f_{LS} = 900$MHz and $f = 1.2$GHz 237
A1 Maximum efficiency for different values of h and σ with 2FRS 247
A2 Maximum efficiency for different values of σ and $h = 0.05L$ with 2FRS 248
A3 Maximum efficiency for different values of σ and $h = 0.4L$ with 2FRS 249
A4 Localization error with 2FRS, $< (5, 4), \mathbf{R}, A >$, $\sigma = 0$ 250
A5 Availability with 2FRS, $< (5, 4), \mathbf{R}, A >$, $\sigma = 0$ 251

A6 Availability on floors 1 and 2 with MSAL-iso, $\sigma = 0$, $h = 0.25L$ 273
A7 Availability on floors 1 and 2 with MSAL-dpl, $\sigma = 0$, $h = 0.25L$ 274
A8 3FRS, $< (5, 3, 3), \mathbf{G}_p, A >$ 276
A9 Percentage of false negatives with $< (5, 3, 3), \mathbf{G}_p, A, iso >$, $\sigma = 0$, $h = 0.5L$ . . 277
A10 Percentage of false negatives with $< (5, 3, 3), \mathbf{G}_p, A, iso >$, $\sigma = 0$, $h = 0.35L$. 278
A11 Minimum false negative probability with $< (5, 3, 3), \mathbf{G}_p, A, iso >$ 280
A12 Minimum false negative probability with $< (5, 3, 3), \mathbf{G}_p, A, dpl >$ 281
A13 Percentage of false negatives with $< (5, 3, 3), \mathbf{G}_p, A, iso/dpl >$, $\sigma = 0$ 282
A14 Probability of false negatives with $< (5, 3, 3), \mathbf{G}_p, A, iso/dpl >$, $\sigma = 5$ 282
A15 Minimum false negative probability with $< (5, 3, 3), \mathbf{G}_p, A, iso/dpl >$, $h = 0.25L$ 283
A16 3FRS Average availability with different decision depths with isotropic antennas, $\sigma = 0$, $h = 0.25L$ 284
A17 3FRS Average availability with different decision depths with dipole antennas, $\sigma = 0$, $h = 0.25L$ 285
A18 Availability on floors 2, 3 and 4 with MSAL-iso, $\sigma = 0$, $h = 0.25L$ 287
A19 Availability on floors 2, 3 and 4 with MSAL-dpl, $\sigma = 0$, $h = 0.25L$ 288
A20 False negatives on floors 2, 3 and 4 with different decision depths with isotropic antennas, $\sigma = 0$, $h = 0.25L$ 289
A21 False negatives on floors 2, 3 and 4 with different decision depths with dipole antennas, $\sigma = 0$, $h = 0.25L$ 291
A22 Average false negatives with different decision depths with isotropic antennas, $\sigma = 0$, $h = 0.25L$ 292
A23 Average false negatives with different decision depths with dipole antennas, $\sigma = 0$, $h = 0.25L$ 293
A24 Localization errors on floors 2, 3 and 4 with MSAL-iso, $\sigma = 0$, $h = 0.25L$ . . 294
A25 Localization errors on floors 2, 3 and 4 with MSAL-dpl, $\sigma = 0$, $h = 0.25L$ . . 295
A26 Minimum localization error with respect to h when $\sigma = 0$, 3FRS 297
A27 Minimum localization error with respect to h when $\sigma = 5$, 3FRS 298
A28 Availability at minimum localization error conditions with respect to h when $\sigma = 5$ 299

A29 Maximum efficiency with respect to h when $\sigma = 0$, 3FRS 300
A30 Average localization error at maximum efficiency conditions with respect to h when $\sigma = 0$ 301
A31 Maximum efficiency with respect to h when $\sigma = 5$, 3FRS 302
A32 Average localization error at maximum efficiency with respect to h when $\sigma = 5$ 303
A33 Maximum efficiency with respect to σ when $h = 0.25L$, 3FRS 304
A34 Average localization error at maximum efficiency conditions with respect to σ when $h = 0.25L$, 3FRS 305
A35 Availability at maximum efficiency conditions with respect to σ when $h = 0.25L$, 3FRS 306
A36 Minimum localization error with respect to h when $\sigma = 0$, 3FRS 307
A37 False negative percentage with respect to h when $\sigma = 0$, 3FRS 308
A38 False negative percentage with respect to h when $\sigma = 0$, 3FRS 308
A39 Minimum localization error with respect to h when $\sigma = 5$, 3FRS 309
A40 False negative percentage with respect to h when $\sigma = 5$, 3FRS 310
A41 Maximum efficiency with respect to h when $\sigma = 0$, 3FRS 311
A42 Maximum efficiency with respect to h when $\sigma = 5$, 3FRS 312
A43 Availability with respect to h when $\sigma = 0$, 3FRS 313
A44 Availability with respect to h when $\sigma = 5$, 3FRS 313
A45 Maximum efficiency with respect to σ when $h = 0.25L$, 3FRS 314
A46 Localization error with respect to σ when $h = 0.25L$, 3FRS 315
A47 False negative percentage with respect to σ when $h = 0.25L$, 3FRS 316
A48 Availability with respect to σ when $h = 0.25L$, 3FRS 317
A49 Performance when optimized for LS, $N = 4$ MoSs, free space, $f_{LS} = 900$MHz, $\sigma = 5$ 318
A50 Performance when optimized for LS, $N = 4$ MoSs, free space, $f_{LS} = 1900$MHz, $\sigma = 5$ 319
A51 Performance when optimized for LS, building height, $f_{LS} = 900$MHz, $\sigma = 5$. 319
A52 Performance when optimized for LS, building height, $f_{LS} = 1900$MHz, $\sigma = 5$ 320

PREFACE

I wish to express my gratitude to everyone who contributed in any way to the completion of this thesis. Among all, I must single out my advisor, Assoc. Prof. Dr. Prashant Krishnamurthy. I appreciate his guidance during the development of this thesis, in addition to his friendship, support and understanding. I would also like to thank Assist. Prof. Joseph Kabara for his guidance during the first two years of my graduate studies.

I would like to thank to the members of my Ph.D. committee, Assoc. Prof. David Tipper, Assoc. Prof. Martin B. H. Weiss, Assist. Prof. Konstantinos Pelechrinis and Assoc. Prof. Nirupama Bulusu for finding time in their busy schedule to review my thesis, to participate in the thesis committee, and for their valuable comments and suggestions to make this thesis better.

I owe thanks to all my classmates and staff in School of Information Sciences for their friendly approach. Very special thanks go out to Assist. Prof. María Calle for her warm friendship making the Pittsburgh cold bearable; to Assist. Prof. Carlos Caicedo for reminding me there are endless opportunities in life once one is ready to embrace them; to my close friend Cihan Ünal for being my roommate, for his understanding in my most desperate times, and for being kind enough to carry all the heavy grocery to home all the time; and to Dr. Zümrüt İmamoğlu for all the tango I danced in Pittsburgh.

I would like to thank to my colleagues and especially my manager at Broadcom Corp., Imran (Hafiz) Hameed, for his encouragement and understanding especially during the last months of my Ph.D. studies.

I am grateful to my uncles, Prof. Dr. Nejat Aksu and Sedat Aksu, who have always been supportive and caring for my entire life. I feel lucky to know that I can benefit from the experiences of these two people with such wonderful characters.

I would like to thank my aunt, Gülten Dikmen, and my grandparents for their love, good wishes and prayers. I feel happy to know that my grandfather and grandmother are in peace and still among us.

I owe a huge debt of gratitude to my parents, Türkan and Talat Aksu and my brother Anıl Aksu for their love and support throughout not only the course of my Ph.D. studies but throughout my entire life. They believed in me more than I believed in myself and this gave me the most strength through years. I miss them, and wish to have more time to spend with them in future.

Finally, I would like to thank my loving husband Melih Korkmaz. I could never even imagine finishing this Ph.D. without his encouragement and endless ways of motivation. His energy and constructive criticism kept me strong, and his compassion relaxed me when I needed the most. His patience and understanding during many stressful and long nights of studying was unfailing and unselfish, and appreciated more than he knows.

1.0 INTRODUCTION

Localization is a term used for determining where a given object or person is physically located in a given environment. For example, the problem of finding the room and the floor on which an asset is located in a building is a localization problem. This problem is also referred to by other names such as location estimation and positioning.

The most mentioned applications of positioning systems are wayfinding, search and rescue, asset tracking, emergency services and smart environments. A localization system can save lives in the aftermath of an earthquake or a snow avalanche; position people or valuable assets during a fire in a building; or track airplanes in the air and at the airport. Other applications are office applications such as nearest printer services; tour and museum guides which can help people navigate in an unfamiliar place; medical facilities that can track staff and monitor patients; and home applications for home entertainment purposes [4]. Also a number of emerging applications exist for localization systems such as monitoring of intruders in wireless ad hoc networks in the case of a security threat or detecting underutilized spectrum with wireless cognitive networks [5].

Wireless networks are becoming pervasive and have been densely deployed for communication of various types of devices such as laptops, mobile phones, handhelds, and sensors. Wireless Local Area Networks (WLANs) are used widely in university campuses and/or libraries to provide Internet connectivity to users. Exploiting these wireless networks for the localization of people or other assets is an added convenience. Especially, the explosion of Wi-Fi, Bluetooth (BT) and other wireless networking technologies has led to many end user devices being equipped with radio frequency (RF) hardware which can be used for location sensing [4].

1.1 CHALLENGES AND MOTIVATIONS

The most popular localization system is the Global Positioning System (GPS) [6, 7] – a worldwide satellite-based radio navigation system based on 24 low-earth orbit satellites. GPS can provide location information for outdoor mobile users with an average positioning error of 7 - 10 meters. The high cost, complexity, and battery consumption in receivers are factors that make GPS undesirable for positioning small sized mobile devices [8]. More significantly, devices that are indoors or blocked by buildings cannot be localized accurately with GPS due to multi-path propagation, insufficient resolution and signal strength in these places.

Especially, the indoor environment is challenging for positioning because of radio channel multi-path effects, signal interference, obstacles, and mobility [9]. To overcome these problems and yet be able to localize indoors, researchers have created several indoor localization systems. These systems may differ from each other in the type of signals (infrared, RF, ultrasound) and the type of signal measurements (RSS, round trip time, etc.) used, and the localization methods utilized for deriving a position from the measurements. Approaches also differ in the assumptions they make about their respective network and device capabilities: device hardware, signal-propagation models, network structure, node density, communication costs, and device mobility [10, 11]. For example, not every wireless device may have GPS chips; and other types of localization that are optimized for suitable signal measurements must be used for these devices. Therefore, employing the right technology and right signal measurements for a given localization application poses another challenge.

Wi-Fi (and the received signal strength – RSS – with Wi-Fi), because of its widespread presence and coverage, has high potential for localization. However, due to the time varying and site-specific nature of the indoor wireless channel, deriving accurate locations from the RSS measurements is not trivial. Such *fine grained* localization systems (using the RSS) rely heavily on cumbersome offline calibration of the features of the RSS for localization [12]. In other words, they need extensive experimental data (obtained prior to system deployment), which is then used to develop the localization algorithm. This process can be time consuming, requires manual labor, and may need to be repeated periodically in dynamically changing

environments. More recently, Wi-Fi localization systems, still using the RSS have been proposed (e.g., [1]), that eschew the costly offline calibration phase at the (presumed) expense of accuracy. They are sometimes called calibration-free localization schemes. *Coarse grained* localization such as proximity based localization does not suffer from multipath and hardware related problems associated with fine grained localization that employs RSS, time, time difference, or angle of arrival measurements. Proximity based localization measures nearness or exploits connectivity information inferred by the quantization of received signal strength. Therefore, it is less complex than fine grained localization. The challenge here is to determine the tradeoffs between accuracy, cost, and practicality of the localization system.

Localization in multi-floor buildings introduces a number of additional challenges. Floor determination problem and location estimation on a given floor must be considered together to determine the parameters that are required for systems with good accuracy and low cost.

1.2 PROBLEM STATEMENT

Localization systems make use of the proximity or distance or directions to monitoring/reference stations with known locations. Although there are quite a number of studies on proximity based localization systems, there is no systematic study available. This is the first problem we address in this dissertation. Our objective is to present *a systematic study of proximity-based localization* by defining and examining the factors and parameters that affect the localization performance. As a result of this study we expect to answer the following question: *is it possible to fine tune these parameters to improve the localization accuracy and what is the accuracy that can be achieved?*

From our efforts to answer these questions, we have determined that the number, monitoring range, and locations of monitoring stations (MoSs) are significant parameters. Basically, the higher the number of monitoring stations is and the more symmetric is the deployment shape, the better is the accuracy. However, a higher number of infrastructure elements will have to be deployed and maintained and this would escalate the overall cost of the localization system. Therefore, an interesting aspect we investigate is the effect of com-

bining the readings from monitoring stations deployed in different floors for better coverage and also better accuracy in each floor of a multi-floor building. The questions we expect to answer here in this dissertation are: *is it possible to utilize monitoring stations across floors to localize the mobile stations with good accuracy, and how much cost and/or time can this approach save us?* Our research helped to determine that the placement of MoSs on multiple floors; the relationship between monitoring ranges of MoSs deployed on adjacent floors; the antenna coverage of MoSs; and the relationship between floor height and monitoring ranges are important parameters to be optimized for the best accuracy that can be achieved with a multi-floor localization algorithm. These parameters can also degrade or improve the performance of localization system in terms of availability. We define availability as the fraction of time/space that a location estimate is available to the user. Then, another question we answer in this dissertation is *for a given multi-floor building, is it possible to adjust these parameters so that the localization performance in terms of accuracy and availability will be improved with a multi-floor localization system?*

In the process of improving the accuracy and availability in multi-floor buildings, determining the correct floor that a mobile station is actually located becomes important. Parameters that are fine tuned for optimizing accuracy may or may not provide a good floor determination performance. On the other hand, when mobile stations cannot be located on the correct floor, the overall availability is negatively affected. Related to these, we want to find answers to the following questions: (1) *How can a floor determination algorithm be incorporated with a multi-floor localization algorithm?* (2) *What aspects of such a floor determination algorithm affect the overall availability of a multi-floor localization system?* (3) *What is the performance of floor determination when the parameters are optimized for good accuracy and availability of multi-floor localization algorithms?*

Proximity based localization studies can be applied to practical problems of interest such as the spatio-temporal spectrum sensing in cognitive wireless networks (CWN). Spectrum sensing is an important problem in CWN, because if it can be realized properly, the expensive and scarce wireless spectrum can be used efficiently by allowing secondary (unlicensed or opportunistic) users to utilize the spectrum holes that are not used by primary (licensed) users. Spatial knowledge of network topology and geometric relations of primary

and secondary users can significantly be of benefit to the detection of spectrum opportunities. Localization algorithms can be employed to estimate the locations of users in primary and secondary networks, and determine available spectrum holes. We want *to demonstrate how proximity based localization can be utilized for sensing transmissions from primary transmitters with the object of determining spectrum holes.* We want to answer to the question: *what could be the effect of a possible frequency offset between the primary network and sensing frequency of the localization system on the performance in localization of primary transmitters and detecting spectrum holes?*

1.3 CONTRIBUTIONS

1. In Chapter II, we provide a brief overview of localization systems, algorithms, comparison metrics, and related issues for completeness. Then, we provide a survey of localization schemes that employ the RSS in infrastructure Wi-Fi networks. We discuss the architectures and algorithms of approaches classified as calibration-based and calibration-free. Calibration-based systems can provide good accuracy given that the channel conditions do not change; but, they require extensive human effort. Lately, researchers have begun to create RSS-based localization systems which are relatively calibration-free and, in effect, shown to be less expensive in terms of the deployment costs and more robust to changes in the wireless medium. It is not straightforward to compare these systems – however, **we have developed metrics that use the data presented in the research related to these systems to compare them quantitatively**.
2. Localization techniques often make use of distances or directions to reference nodes with known locations. Such schemes assume *range based localization* using sophisticated measurements such as time or angle of arrival (ToA/AoA) to estimate a node's location [13]-[14]. An alternative approach is *proximity based approach* using only the connectivity information to reference nodes. This approach may have larger granularity than range based approaches depending on the number of infrastructure elements utilized per unit area; however, it does not suffer from multi-path and hardware related problems asso-

ciated with range-based localization. In Chapter III, **we propose a proximity-based localization algorithm, *Sub-Area Localization* (SAL), that uses the sub-areas created by the overlapping ranges of monitoring stations (MoSs) to estimate the location of a mobile node.** We investigate the relationship between localization accuracy (that depends on the number and sizes of sub-areas) and the monitoring ranges of MoSs when they are placed on a virtual grid in a given workspace. We present analytical and simulation results that allow us to determine the "best" range of monitoring stations, and understand the limits on the accuracy and availability performance of SAL. We also investigate the channel effects such as log-normal shadowing, and environmental effects such as wall attenuation factor. We study the effect of MoS placement schemes on accuracy of SAL through simulations of deviation from grid placement, random placement and incremental placement of MoSs.

3. Proximity based localization schemes in the literature have predominantly focused on single-floor or 2-D localization. Multi-floored environments provide us an opportunity to exploit the coverages of MoSs that are deployed across floors. In Chapter 4, we focus on the **possibility and advantages of utilizing MoSs across floors if they can sense transmissions across floors to achieve a satisfying accuracy with reduced deployment cost.** We call this proximity-based multi-floor localization approach as Multi-Floor Sub-Area Localization (MSAL). We first present the performance of MSAL when floor information of MNs are known through an external floor determination mechanism. This could be possible through requesting an input from the user, or keeping track of the floor the MNs are on through RFID tags placed at each floor of the multi-floor building. Under this assumption, we investigate the effects of various parameters such as the relationship between monitoring ranges of MoSs on different floors, the relationship between the floor height and monitoring ranges, MoS placement schemes and the antenna coverage of MoSs on the accuracy and efficiency of a the localization system. Our efficiency definition comprises good accuracy by utilizing as few MoSs as possible per floor in a multi-floor building. We study the effect of log-normal shadowing, and other signal propagation models based on wall attenuation factor and floor attenuation factor through simulations. Because placing MoSs on exactly the same locations on

different floors may not be possible, we investigate the effect of placing MoSs on random locations on each floor in a multi-floor building. Then, we remove the assumption of known floor information, and propose that a majority based floor determination can be integrated with MSAL. We study the effect of floor determination mechanism on the availability and efficiency of localization system, and the tradeoffs between floor determination mechanism and MSAL. Although it is not possible to change the floor height in a given multi-floor building, our findings imply that there is an optimum floor height to satisfy a given accuracy level.

4. Localization algorithms can be employed for spatio-temporal sensing in wireless cognitive networks. Spatial knowledge of on network topology can be used to identify and utilize the underutilized spectrum. In Chapter 5, **we demonstrate how Sub-Area Localization can be utilized for spatio-temporal sensing of primary transmitters with the object of determining spectrum holes.** The network may include primary transmitters operating on different frequencies. Localization system can monitor these transmissions, however cannot modify the best monitoring range for each different frequency once it is deployed for service. When there is an offset between primary users' frequency and the frequency for which the localization system is optimized, the estimation accuracy will get affected. Therefore, we also evaluate the performance of SAL in localization of primary transmitters and detection of spectrum holes when there is a frequency offset between the localization system and primary network. Finally, we provide discussion on how SAL's performance can be improved with different scenarios such as using secondary network topology information for spectrum hole detection, or the effect of the location knowledge of primary receivers, etc.

2.0 LITERATURE REVIEW

2.1 LOCALIZATION SYSTEMS AND PERFORMANCE METRICS

Localization techniques differ in the assumptions they make about their respective network and device capabilities: device hardware, signal-propagation models, network structure, node density, communication costs, and device mobility [10, 11]. In order to measure the performance of these techniques various metrics can be used. These metrics are accuracy, precision, robustness, adaptation, coverage, scalability, responsiveness, availability, cost, endurance to security attacks, and efficiency in deployment.

The coordinate systems employed by localization schemes can be local, relative, or absolute [15]. An absolute coordinate system (e.g., used in GPS) has *global coherence* and is aligned with the coordinate systems used in commercial and military applications. In a local coordinate system, the communicating devices position themselves only with respect to each other. The estimated positions are relative to a local system, can be arbitrary, but may still provide network-wide coherence. In these systems, the location coordinates are calculated with respect to the coordinates of several predefined and fixed *reference points* (RPs). An RP can be a cellular base station, a Wi-Fi access point (AP) or a USB Bluetooth (BT) adapter. The number and locations of RPs have significant impact on the performance of localization systems. Therefore, localization methods must be *robust* against RP failures or non-uniform RP distribution [16].

The deployment of location estimation systems can be classified into two types: client-based (distributed) and infrastructure-based (centralized). In client-based systems (e.g., GPS), also called self-positioning, location estimation is done with the information collected at client devices; in infrastructure-based systems, also called remote positioning, localization

depends on the information collected at RPs and the location is calculated at a server.

The *accuracy* of a system denotes how close the estimated location is to the actual location. The *precision* is with what probability we can get that accuracy [10]. For example, some of GPS receivers can reach 1 to 3 meter accuracy 99 percent of the time. When the object of interest is static, localization algorithms are used to identify its location; however, when the object of interest is moving, tracking algorithms are used to estimate its *path* over time. Localization algorithms or tracking algorithms can be used for one another in order to achieve better accuracy. For example, if the localization error is larger than the maximum distance that a mobile can travel with a given speed in a given time duration, then the mobile's previous location can be used to limit the distance between the previous and present location estimates. However, this may cause error propagation. Our focus is on the localization algorithms rather than tracking; interested reader may refer to this paper for more details [17].

The *adaptation* of localization systems in the presence of changing environmental conditions is also required for good performance. Localization systems must be adaptive to noisy environments, or should be able to provide the required accuracy and precision in different environments with different conditions.

A localization system must be able to locate objects in a given environment in order to satisfy the requirements of the target application. This environment may be the world, a campus, a building or a room. However, whatever the *coverage* is, the number of objects a system can locate with a certain amount of infrastructure or over a given time may be limited. For example, while a master BT device can connect up to 7 slave BT devices, a tag reader may not read any tag if more than one tag is within the range [10]. The coverage area per unit of infrastructure, and the number of objects the system can locate per unit of infrastructure per time interval values are used to determine the *scalability* of a localization system [10]. Therefore, in order to provide a good scalability, the locations must be estimated within a fairly low response time with given infrastructure. The *responsiveness* of the system is determined by the latency in estimating the locations of objects [10], and it is affected by factors such as available bandwidth, and the computational overhead of the technique. The *availability* is the fraction of time/space that a location estimate is available

to the user. Availability may depend on coverage, channel propagation characteristics and location estimation algorithm capabilities.

The *installation cost* of a system includes the cost due to hardware and software requirements, and the wages of people who installs the system, and the *maintenance cost* includes the wages of people who fix devices not functioning [15].

Localization techniques are highly vulnerable to *security* attacks from dishonest nodes and external attackers. Dishonest nodes can report false positions and distance information in order to cheat on their locations, whereas external attackers can spoof measured locations of honest nodes. An attacker can influence localization measurements by jamming and delaying signals and by modifying their signal strengths besides setting wormhole attacks or cloning attacks. A number of secure localization techniques which use multiple kinds of measurements [18], distance-bounding protocols [19] or verifiable multi-lateration [20] were proposed in the literature. Localization techniques considered in the following sections assume non-adversarial settings.

2.2 LOCALIZATION ALGORITHMS

Localization systems rely on some form of communication between RPs with known locations and the object that needs to be localized. Depending on the granularity of information inferred by this communication, indoor localization approaches can be classified into two categories: coarse grained and fine grained [16].

2.2.1 Coarse Grained Localization (Proximity-based)

Approaches that infer coarse grained information such as proximity based approaches fall into the coarse grained localization category. Proximity based localization measures nearness or exploits connectivity information to a known set of RPs [21]. This approach is also called cell-id approach. Connectivity is inferred by the quantization of received signal strength at the RP. If the received signal strength is higher than or equal to reception sensitivity, then,

mobile device is said to be connected to RP. Figure 2.1 shows two examples of proximity based localization. In the first example (upper figure) the mobile device at point P is only connected to RP P_3, then, the location is approximated as the location of the RP P_3. In the second example (lower figure), the mobile at point P is connected to three RPs, then, the estimated location is the centroid of the triangle formed by three RPs, P_1, P_2 and P_3. This method is called centroid algorithm [16]. According to this algorithm, the estimated location is calculated from,

$$P(x,y) = \left(\frac{\sum_{i=1}^{N} x_i c_i}{\sum_{i=1}^{N} c_i}, \frac{\sum_{i=1}^{N} y_i c_i}{\sum_{i=1}^{N} c_i} \right), \tag{2.1}$$

$$c_i = \begin{cases} 0 & P_r < R_s, \\ 1 & P_r \geq R_s \end{cases}$$

where c_i is 1 if mobile node is connected to RP P_i, 0 otherwise. (x_i, y_i) denotes the two dimensional coordinates of RP P_i, $i = 1, 2, \ldots, N$.

Electronic article surveillance (EAS) systems that are used in retail and library settings are an example to this approach. The EAS tags are very simple; they deter amateur theft by an alarm which goes off when a resonance due to the tags response to a matched electronic field. Although these systems are low cost, they have a restricted range, limited reliability and they lack identification codes [22].

Another example is The Active Badge [23] system which is an IR-based localization system. In Active Badge, IR identification (IRID) tags periodically transmit their identification codes by emitting infrared light to readers installed throughout the facility. Although IR can provide accurate location information due to its short range, IR-based localization systems suffer from poor scalability due to the same reason. In addition, IR-based systems have high implementation and maintenance cost, since tag prices are relatively high, and a special hardware has to be installed to the mobile device. These systems also perform poorly in the presence of direct sunlight and hence cannot be used outdoors [16].

Landmarc [24] is a Radio Frequency Identification (RFID) based localization system. RFID tags are detectable up to about 3 meters away as they pass fixed readers. Inexpensive RFID tags can be used for location determination by placing RFID readers at doorways

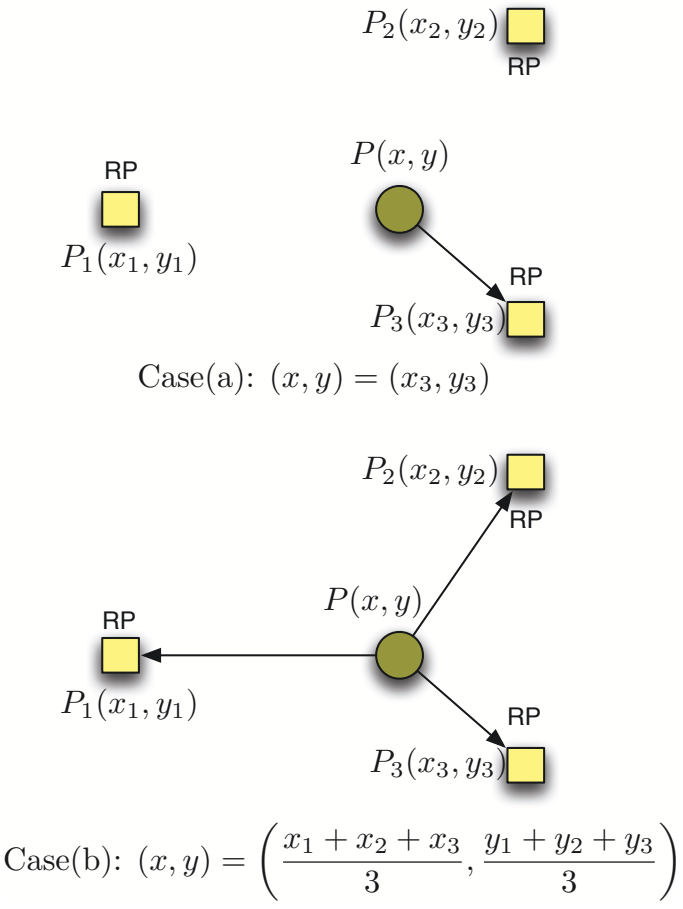

Figure 2.1: Proximity-based localization

and other strategic points to detect the passage of people or objects [4]. This causes poor scalability. RFID systems are being used for handling automated toll collection on highways.

In [25], a proximity based localization algorithm for large scale sensor networks (APIT – approximate point-in-triangulation test) is proposed. APIT uses RPs to perform location estimation by isolating the environment into triangular regions between RPs. A node's presence inside or outside of these triangular regions allows a node to narrow down the area in which it can potentially reside. This algorithm utilizes the combination of RPs which results in the highest number of triangles intersecting at the location of the sensor node. The center of gravity (COG) of the intersection of all of the triangles in which a node resides is used to determine the estimated position. This algorithm is distributed, and sensors also

utilize neighbor information to decide whether they are inside or outside a triangle. The simulation results show that the performance of APIT depends on the node degree (a node degree of 8 is used in [25]). When RPs are distributed uniformly in a grid, at least 10 RPs must be heard by a node to achieve a localization error of 70 % of the radio range. For most of the experiments in this paper, 16 RPs are heard by localizing nodes.

Computer vision based systems use cameras and image processing techniques for localization. Their drawbacks are the difficulty identifying many subjects simultaneously, high cost and poor scalability.

In general the advantages of proximity based localization are good responsiveness due to short computation times, and practical solution as collecting measurement metrics for more fine-grained localization is quite complex. The major drawback is that the accuracy depends highly on the number and locations of RPs; therefore, scalability may be a problem when good accuracy is required.

2.2.2 Fine-Grained Localization

Approaches that infer fine grained information such as distance-based, angle-based and scene analysis approaches fall into the fine grained localization category.

2.2.2.1 Distance-Based Approach: In distance-based methods, the distances of the node that needs to be located to several RPs are determined by measuring one of the metrics such as Time of Flight (ToF)/ Time of Arrival (ToA), Time Difference of Arrival (TDoA), Received Signal Strength (RSS) or Signal-to-Noise Ratio (SNR). Depending on the deployment of the system, these metrics can be measured at the RPs and sent to a central server to be used in the location estimation algorithm (infrastructure-based), or can be measured at the mobile node and used in the location estimation algorithm applied at the mobile (client-based). Then, the position of the node is estimated using the *lateration* technique.

Lateration is a commonly used location estimation technique in which the distance measurements are acquired from $(n + 1)$ non-collinear RPs in order to estimate the location of an object in n-dimensional space [11]. In Figure 2.2(a), let P be the mobile to be located

in a 2-dimensional plane. Point P known to be at a distance d_k from point P_k must be located on a circle of radius d_k centered at point P_k, where $k = 1, 2, 3$. Three such circles around RPs placed in 2-dimensional space, such that they are not collinear, will intersect at one point. Point P must be at this intersection point in order to generate the observed distances. Then, the estimated location is found from the following set of equations, since there are 3 equations and 2 unknowns, (x, y):

$$d_i = \sqrt{(x - x_i)^2 + (y - y_i)^2},\ i = 1, 2, 3, \tag{2.2}$$

where (x_i, y_i) are the coordinates of RP_i.

When errors at distance estimations occur, these circles do not intersect or they intersect to define an area rather than a single point as shown in Figure 2.2(b). In this case, more complex estimation algorithms are needed to locate P.

In timing based approach, the distance from a point P to a RP is inferred from the Time of Flight (ToF) or Time of Arrival (ToA) of the communication signal [16]. Measuring ToF means measuring the time it takes a signal to travel between the point P and a RP at a known velocity. The basic equation for distance is the multiplication of velocity and the time it takes the signal to arrive. The transmitted signal carries a time stamp showing when it is transmitted. When the receiver receives the signal, it calculates the current time, and the difference between the current time and the time stamp at the signal gives ToF.

Ultrasound signals can be used to determine distances between mobile tags and known points in the environment. Sound waves have a velocity of approximately 344 meters per second in 21°C air [21]. Then, an ultrasound pulse sent by an object and arriving at point P 14.5ms later implies that the RP is 5m away from point P. Active Bat [13] localization system uses an ultrasound ToF lateration technique to provide accurate positioning. Because ultrasound signals have short wavelengths, location accuracies on the order of centimeters can be reached. A major problem is that the propagation velocity of the ultrasound is easily affected by the temperature and humidity, which introduce ranging errors [26].

A light pulse emitted by an object has a velocity of approximately 300×10^6 meters per second, and will travel 5 meters to a point P in 16.7ns. Radio waves travel with the same

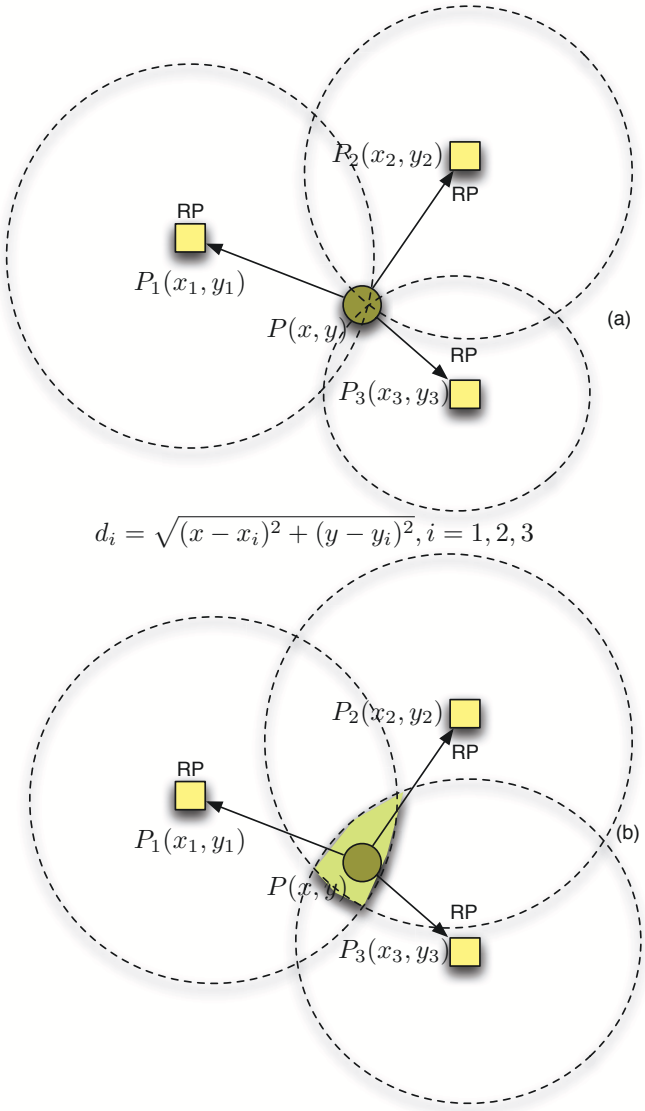

Figure 2.2: Lateration (a) with no errors, (b) with errors

velocity as light. Therefore, measuring ToF with radio is possible but requires clocks with much higher resolution than those used for timing ultrasound.

Pinpoint's 3D-iD [27] uses base stations and tag hardware to measure RF ToF. It uses an installed array of antennas at known positions to perform lateration. Pinpoint's accuracy is roughly 1 to 3 meters. Cricket location system [28] uses both RF and ultrasound signals to measure the ToF, and the distance between the RP and the object to be located is found

based on the difference of arrival times of these signals. In Ubisense Localization System [29], ToA measurements were used with UWB technology. Unlike conventional radio signals, UWB signals can have pulse durations short enough to allow accurate ToA with an accuracy of 15 centimeters. The precision capabilities combined with the very low power makes this approach ideal for certain radio frequency sensitive environments such as hospitals and health care. Also, UWB technology does not require a direct line of sight between tags and readers; however, it has a high cost of initial implementation [4].

ToA method requires accurate synchronization between the RPs, but the node itself might have a drift that can reach a few microseconds; this drift generates an error in the location estimate. TDoA is used to help avoid node clock synchronization errors [30]. HP Labs SmartLOCUS [31] uses synchronized RF and ultrasound differential time-of-flight measurements to determine the inter-nodal range between any two nodes. This system yields an accuracy of 2-15 cm.

Moreover, indoor environments cause a lot of reflections and multi-paths; therefore, it is very important to calculate the ToF of the first arriving ray which is assumed to be coming through the most direct path between a RP and a wireless device. The second and third components of the signal follow the first ray, probably after bouncing of off doors, walls, and other furniture. If the delay between the first ray and the others is smaller than the duration of the pulse shape used in the wireless system, these two rays overlap, causing errors in ToA [30]. Therefore, line of sight (LOS) is an important requirement for high accuracy in timing based localization systems. If the environment is likely to have so much non-LOS (NLOS), then the number of infrastructure can be increased at the expense of cost in order to increase the probability of LOS and, in effect, to provide more accurate results.

According to the Friis free space radio model, the intensity of a signal emitted by an object is attenuated by a factor proportional to $1/r^2$ when it reaches a point P at distance r from an object [32]. The intensity of the signal is measured as RSS at the receiver and is reported as the Received Signal Strength Indication (RSSI) in dB. According to this model, RSS is a function of signal's transmission power, P_t, the distance between the receiver and the transmitter, r, the path loss of the environment, n, noise, N_0 and the channel coefficients,

h [33].

$$RSS = log_{10}\left(\frac{P_t \|h\|^2}{r^n N_0}\right)$$

In Signal Strength Based approach, knowing all the other parameters, one can calculate the distance r from an AP to a wireless device. However, multi-path reflections, shadowing effects, diffraction around sharp corners or scattering from wall, ceiling, or floor surfaces, and absorption by structures and even human bodies make it difficult to accurately characterize the large-scale (signal propagation) and small-scale (fading) effects on signals inside a building using a basic model, and also degrade the link performance. For these reasons, in realistic situations, deriving location from RSS is not trivial. The effect of reflecting and attenuating objects in the environment can have much larger effects on RSS than distance, making it difficult to infer distance from RSS without a detailed model of the physical environment. Two wireless devices that are apart from an AP by the same distance can have quite different RSS measurements due to different intermediate obstacles between the devices and the AP [26].

It has been shown that increasing the transmission power does not help to improve the localization accuracy, but rather degrades the performance [34]. This is due to the fact that the increase in the transmission power makes the wireless channel more noisy as a result of multipath fading and results in poor estimation of signal strengths.

The preprocessing of RSS samples is the first alternative proposed against the unpredictable nature of RSS. It has been shown that a median filter can be used to deal with short-term variations in RSS due to fast fading [3]. A median filter takes the median value of RSS samples measured for a certain time duration.

Various signal propagation models are utilized for more accurate distance prediction in environments with different characteristics. An example is Wall Attenuation Factor (WAF) based signal propagation model [35]. In this model, the signal path loss is predicted through counting the number of obstructions between the transmitter and the receiver. Path loss in dB that is caused by one obtruction is denoted as WAF, wall attenuation factor. This

model is described by,

$$P(d) = \begin{cases} P(d_0) - 10nlog\left(\frac{d}{d_0}\right) - nW * WAF & nW < C, \\ P(d_0) - 10nlog\left(\frac{d}{d_0}\right) - C * WAF & nW \geq C \end{cases}$$

where n is the rate at which the path loss increases with distance, $P(d_0)$ is the signal power at some reference distance d_0 and d is the distance between the transmitter and the receiver. C is the maximum number of obstructions(walls) up to which the attenuation factor makes a difference, nW is the number of obstructions between the transmitter and the receiver. The values of n and WAF depend on the building layout and construction material, and are derived empirically. The value of $P(d_0)$ is also derived empirically. This model is used to compute the distance between an object and the RP from RSS samples. The distance information is then used to locate the object by lateration. RADAR localization system utilizes this model [3].

When compared to timing based systems, RSS measurements are affected more significantly by fast fading or the environmental changes such as temperature and humidity. However, RSS based systems do not require additional hardware to provide the synchronization and a fine clock; and therefore, are cheaper than timing based systems. Since no additional hardware is required, WLANs that are already deployed can easily be utilized with no additional cost in RSS-based systems. Therefore, widespread deployments of the 802.11 infrastructure and the ubiquity of Wi-Fi embedded devices is a compelling reason for RF based indoor localization. Using RSS, it is possible to locate Wi-Fi enabled devices with accuracies from several meters to tens of meters [3].

Bluetooth technology [36], which can offer shorter range (BT class 2 range is approximately 10 meters) than Wi-Fi, can provide more accurate positioning, but requires more fixed APs to provide coverage. Off-the-shelf BT adapters and devices cannot be used for ToA measurements since they do not have the required fine clock, but can be used with RSS easily. In [37], inexpensive BT USB adapters and BT identification badges are used with RSS measurements.

Signal-to-Noise Ratio (SNR) can also be used as a measurement metric but it has been shown that signal strength was more indicative of location than SNR [3].

DC electromagnetic field based technology can be used in localization and tracking [21] especially in the computer animation industry. Although it can provide good accuracy, it is quite expensive and, like IR, severely range limited; hence, unsuitable for large-scale deployment.

In [38], the ambiance of a place is used to define the logical location of a mobile. The hypothesis of this work is that the combined effect of ambient sound, light, color of a place and the user movement can be used in a localization algorithm. The results from 51 different stores show that this technique can achieve an average accuracy of 87% when all sensing modalities (optical, acoustic and motion) along with RSS are employed.

2.2.2.2 Angle-based approach: Angle-based localization systems use *angulation* technique to estimate the location of an unknown point. In this technique, Angle of Arrival (AoA) measurement which shows from which angle the signal has arrived to a receiver is used. For two-dimensional angulation, two angle measurements and one length measurement such as the distance between the RPs are needed. For three-dimensional angulation, one length measurement, one azimuth measurement and two angle measurements are needed [21]. When the distance between the RPs is not available, triangulation method shown in Figure 2.3 can be used for positioning. A constant reference vector as 0° can be used for angle measurements. In order to measure the AoA, phased antenna arrays, beamforming or directional antennas can be used. The orientation of antenna array affects the performance; therefore, some calibration may be needed [3].

The VHF Omni-directional Ranging (VOR) technique used in aircraft navigation systems is an example of angulation. In [14], authors use the idea of VOR for indoor positioning and using WLAN APs with revolving antennas. In [39], authors combine AoA and RSS measurements and show that the use of directional antennas averages out multi-path effects in indoor environments, and also helps reduce the amount of data needed.

An advantage of using AoA is that synchronization required in time-based measurements is not required in this case [30]. However, the hardware needed to measure AoA has high cost. In addition, as in the time-based systems, measuring the first arriving ray is important, since it will provide the most accurate angle measurement. Therefore, LOS is also a requirement

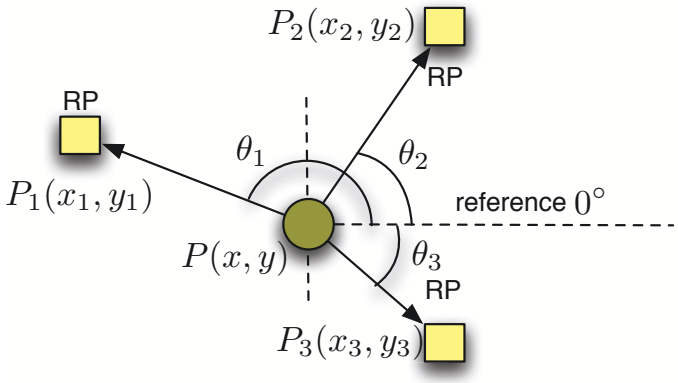

Figure 2.3: Triangulation

for accurate angle-based localization.

2.2.2.3 Static Scene Analysis Approach: Static scene analysis (SSA) is an empirical technique which uses the features of a scene observed from a particular vantage point to estimate the location of objects in the scene [21]. The scene can consist of visual images (frames captured by a camera) or a measurable physical phenomenon (electromagnetic characteristics that occur when an object is at a particular position and orientation).

With SSA, objects can be localized using passive measurements that do not correspond to distances. SSA has two phases. In the first phase, a dataset is obtained by observing the scene. This phase is an offline learning or calibration phase, where the samples (RSSI, ToA, etc.) from known locations are collected and processed so that the relationship between the location and the measurement metrics can be revealed. The second phase is the online localization phase where the location is estimated by matching the measurements collected in real time to the data collected in the calibration phase. The accuracy of estimation depends on the technique used to construct the database, and the algorithm used to match the measured signal to the database. This method is used in RADAR localization system [3], which will be explained in detail in the next sections.

Although SSA eliminates the need to convert the measurements to distances, it requires

extensive calibration in order to create the dataset. Also, changes to the environment in a way that changes the observed features of scenes may require regenerating the predefined dataset.

In the literature, there are attempts to determine the Cramer-Rao Bound (CRB) on the *variance of the localization error* of proximity based and distance based localization schemes in sensor networks. In [40], the distribution of errors in measurements of RSS, ToA, and AoA are used to calculate the CRB for cooperative and range-based localization in sensor networks. In [41], the CRB is calculated for distance-free localization algorithms where sensors only use their neighbors to estimate their own locations. If Y is the maximum distance that a node i can move in any direction without changing its neighborhood, then Y is the resolution limit for the localization algorithm. When the neighborhood does not change, the proximity information stays the same. Y is a random variable and $E[Y]$ is the maximum achievable accuracy of the localization algorithm.

2.3 RSS BASED LOCALIZATION

As discussed previously, the usage of RSS to localize a mobile device in indoor areas is attractive because there is no need for additional hardware in the mobile device nor extra spectrum for localization purposes. In what follows, RSS based localization schemes are classified into two groups as calibration-based and calibration-free localization schemes.

2.3.1 Calibration-based Localization

2.3.1.1 Why Calibration? The RSS measured by a WiFi enabled device is not constant at a given location. Experiments have showed that RSS measurements at a fixed position varies over time and the variations can be as large as 10 dBm [42]. The time variation of the channel can be due to changes in the physical environment such as movement of people [32]. A study of the properties of the RSS, specifically for localization purposes is presented in [43]. The RSS values change with orientation of the user holding the mobile

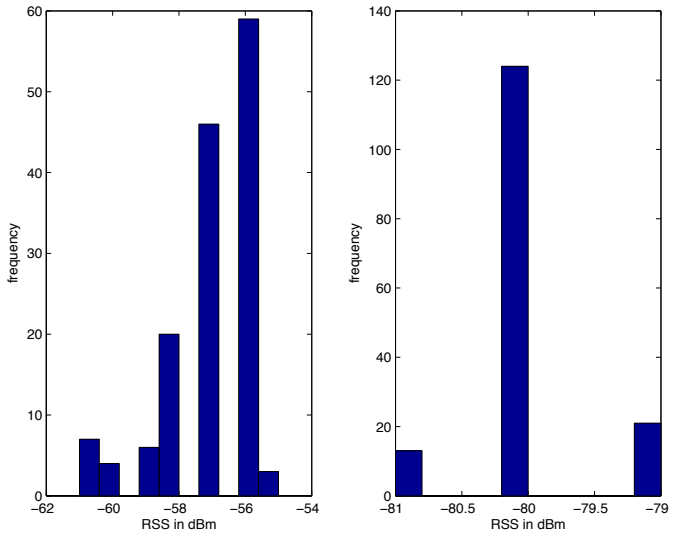

Figure 2.4: Histograms of the RSS from an AP, observed at a mobile station (a) closer to the AP and (b) away from the AP

device and are sensitive to the chipsets used in the WiFi card. It is also difficult to use standard path-loss models since even the mean path-loss is site specific and depends on the floor plan, construction materials, and other objects in the environment. If samples of the RSS are collected over windows of time at a fixed location and analyzed, it is observed that the histogram has a left skew closer to the transmitter (access point) because of limits to the transmit power while the histogram is closer to a symmetric normal distribution at larger distances from an AP as shown in Figure 2.4. Further, the RSS is non-stationary and can see jumps in the mean value over time. In any case, depending on a single sample for estimating the location of the device may lead to large localization errors, especially if this sample comes from the tail of the distribution. Consequently, modeling the behavior of RSS samples prior to location estimation can mitigate the problems due to the dynamic characteristic of the wireless channel. For instance, calibration-based techniques can employ more than one sample of the RSS in estimating the user's location. During calibration, information presented by weak signals or the absence of signals at a location may be characterized.

2.3.1.2 Localization Technique Calibration-based localization systems consist of two phases: calibration phase and location estimation phase. Calibration is an offline learning phase; and location estimation is an online phase, where the actual location finding is realized.

In the first phase, the purpose is to relate the measurement values to the coordinates of the known locations. This relationship can be revealed with either parametric or non-parametric approach. In parametric approach, propagation models are used. These models take into account several variables like the type and the number of walls and other signal obstructers, e.g. WAF based signal propagation model used in [3]. In the non-parametric approach, calibration, the measurement or sampling of available metrics at known locations, is performed. Studies suggest that propagation models are not competitive against empirical calibration models in terms of positioning accuracy due to the insufficient precision of signal models [3].

Modeling the behavior of RSS samples prior to location estimation can mitigate the problems due to the dynamic characteristic of the wireless channel. This can be shown as the most convincing motivation to calibration efforts. During calibration, information presented by weak signals or the absence of signals at a location is also significant. For instance, while a high RSS can imply a short distance between the location and the RP, a low RSS or the absence of RSS can imply a long distance.

The output of the calibration phase may differ from system to system. It can be a radio map containing all measured metrics from all APs, at all possible known locations; or it can be a trained neural network containing all the weights of the connections between the layers and units of the network. In addition, instead of using the raw data gathered, employing data processing or filtering may facilitate the location estimation algorithm in the second phase. For example, the output can be a table for each RP containing the mean and the standard deviation of the RSSs corresponding to each combination of location and orientation of antenna [3].

The effectiveness of calibration depends on two types of parameters. The first type has a unit per training (calibration) location, e.g. the number of samples per location, the sampling interval per location, the number of orientations of antenna per location. The second type is for the whole calibration area, e.g., the number of training locations, the density of the

training locations, times of day and days of week that the measurements are collected.

In location estimation phase, RSS values at an unknown location are measured. Using this run-time RSS, the location database created at the calibration phase is searched for the closest match among the RSSIs of the known locations using a localization algorithm. Note that the bigger the calibration dataset is, the worse the responsiveness of the system may be. Therefore, efficient search algorithms must be employed in location estimation phase.

In the following, various calibration based localization algorithms are discussed in two groups: deterministic and probabilistic. Probabilistic techniques calculate for each known location the probability that the device is at that location. Then, the location with the highest probability is considered as the estimated location of the device with that probability. Deterministic techniques estimate the location without any probabilistic claim.

2.3.1.3 Deterministic Calibration-based Approaches

Empirical Trilateration (ET) uses calibration data in order to estimate the distances to RPs from an unknown point instead of using signal propagation models as in regular trilateration. Thus, the radius of each circle centered at RPs is determined by matching the run-time RSS values to those of the offline collected at the calibration phase. The centroid of the smallest perimeter triangle created by any three RPs is the estimated location. When compared to traditional lateration, ET improves the accuracy by 11.8% [44].

Smallest M-vertex polygon (SMP) uses the run-time RSS values to search for a number of candidate locations in signal space with respect to each RP separately [44]. M-vertex polygons are formed by choosing at least one candidate from each RP (total of m RPs). SMP chooses the smallest polygon, which has the shortest perimeter. Averaging the coordinates of vertices of the smallest polygon gives the final location estimate. When compared, SMP and ET have similar mean errors; however, the variation of error of SMP is less than the variation of error of ET.

Fingerprinting is proposed as another deterministic calibration based localization approach in [45, 3, 46, 47]. This method is an example of SSA explained in previous sections.

In the first phase of fingerprinting, the entire area is covered by a rectangular grid of points. The RSS is measured with enough statistics to create a database of predetermined

RSS values on the points of the grid. The vector of RSS values at a point on the grid is called the *location fingerprint* of that point. Each entry in the database includes a mapping of the grid coordinate $l = (x, y)$ to the vector of corresponding RSS values from all RPs in the area, $\mathbf{s}_l =< RSS_{l1}, RSS_{l2}, \ldots, RSS_{lm} >$, when there are m RPs. This process is also called RF profiling or RF mapping.

In the second phase, a sample of the RSS from all RPs at the user's current position is obtained. This run-time RSS vector is compared with all existing entries in the database. The fingerprint entry that has the closest match to the user's sample of RSS is returned as the estimate of the user's current location.

In [45], each element in each vector in the database is assumed to be the mean of the RSS from each of the RPs in the area. *Least squares error (LSE) algorithm* estimates the location by computing the Euclidean distance between the run-time RSS vector, RSS, and each fingerprint in the database. The coordinates associated with the fingerprint, $l = (x, y)$, that provide the smallest Euclidean distance are returned as the estimate of the position,

$$(x, y) = argmin_{l \in L} \sum_{i=1}^{m} (RSS_{li} - RSS_i)^2$$

RADAR [3] system considers also the orientation of the antenna on the training locations. For each location-orientation tuple, (x, y, θ), the standard deviation, and the median of the corresponding RSS values are computed for each of the RPs. The processed data set rather than the original raw data set is used in the second phase. RADAR uses *Nearest Neighbor in Signal Space (NNSS)* algorithm to find the closest match. The location with the minimum Euclidean distance between the real-time data and the database is returned as the estimated location.

Another algorithm which is similar to NNSS is *K-nearest neighbor averaging (KNN) algorithm*. KNN searches for K closest matches by comparing the root mean square (rms) errors between the run-time and offline RSS vectors. Averaging the coordinates of the K-locations gives the final location estimate. In [1] a performance comparison of ET, KNN and SMP was held, and it was shown that KNN performs better than ET and SMP yielding the smallest mean distance errors.

In traditional fingerprinting approaches, in order to create the database, RSS samples are collected from known locations, which usually are the vertices of a virtual grid at the experiment area. LEASE system [46], on the other hand, instead of using training locations, uses stationary emitters (SEs) in addition to sniffers (APs). The SEs are wireless transmitters that send a few packets occasionally, and they *are not* necessarily placed to form a grid. The coordinates of the SEs are known to the location estimation engine (LEE). The sniffers sniff on the wireless medium, and listen for all communication from wireless clients and SEs. They behave like monitors instead of reference points.

In calibration phase, the LEE collects from all sniffers the RSS from the SEs, and uses this information to build its RSS model for location estimation. The RSS is modeled for each sniffer as a function of the coordinates of the SEs at a site. First, the data points are smoothed using a generalized additive model (GAM) [46]. Second, the site is divided into small grids and using Akima splines, the smoothed values obtained from the GAM are interpolated to estimate the RSS at each grid center. To generate this synthetic model of RSS grid, only the coordinates of the SEs and the RSS from the SEs are needed. This technique is repeated for each sniffer deployed, and at the end of this process, a set of grids for the site is obtained, where each grid has an associated n-vector of estimated RSS, when there are n sniffers on the site. Calibration in LEASE is not limited by the offline phase, it is continuously processed also in online phase. Assuming RSS follows a log-normal distribution, the system rebuilds the model when a statistically significant deviation in RSS coming from an SE is observed by at least one sniffer. With this approach, a model will be rebuilt if an SE or sniffer is moved, or the environment changes significantly.

As a difference from other proposed works, LEASE justifies its approach with the experiments done in two different floors in a site. In order to locate a client, LEE uses the RSS information from the client as recorded by all the sniffers, first to map the client to a floor at a site, then, to estimate the client's location by "matching" the run-time and offline RSS vectors. Mapping a client to a floor is realized by sorting the RSS coming from sniffers and applying the majority rule. According to the majority rule, the floor which has the majority of the sniffers from which m strongest signals are seen is the floor where the client is located. After the floor is estimated, the client's RSS vector can be matched to that floor's model by

using *Nearest Neighbor Search (NNS)* algorithms. Two variations of NNS are used in this work: *Full-NNS*, and *Top(k)-NNS*. Full-NNS matches the entire RSS vector as seen from the client to the RSS vectors at each synthetic grid point to find the closest match. However, Top(k)-NNS considers only the top k RSS from the client, and matches the client's RSS k-subvector with the corresponding k-subvector of only those grid points where the sniffers with the top k RSSs.

[46] investigates the effect of the channel sampling interval, the number of sniffers and SEs. Minimizing the number of SEs was shown to be equivalent to minimizing the amount of profiling needed. Another work which investigates the effect of reducing manual calibration on the accuracy of the localization system is [47]. This work uses the ***interpolation of RSS*** calibration data taken from known locations. It is shown that it is unnecessary to spend much time at each location, as more time beyond a threshold does not improve accuracy very much.

Experiments showed that fingerprinting approaches yielded higher accuracies than the approaches using signal propagation models; however, they require extensive calibration effort.

[37], [48], [49, 50, 51, 52, 53] propose ***Neural Networks (NN)*** as alternative to previously mentioned calibration based localization methods. A neural network is an interconnected group of artificial neurons, and it uses a mathematical or computational model for information processing [54]. NNs differ from other algorithms in that they learn from experience allowing them to improve their performance and adapt themselves to changes in the environment.

In the calibration phase, RSS measurements collected from multiple APs are fed into a neural network, and this neural network is trained so that the distance between the output location and the target location is minimized. At the end of the calibration phase, the neural network has all the weights and parameters calculated. At the location estimation phase, the system collects run-time RSS and feeds it into the neural network to get the output as the estimated location. There is no need to interpolate the training locations, since this is already done by the neural network itself [37].

A number of studies exploiting neural networks for location estimation or location track-

ing were published. These works differ from each other in terms of the location measurement methods, the types of neural networks, and training functions used. Among these studies [49, 50, 51, 52] use RSS measurements as input to the neural networks, whereas [53] uses ToA. For instance, [48] feeds RSS measurements from 5 different APs to the neural network, and the output is the two dimensional coordinates of the device that is to be located. The most important contribution of this paper is that it contains the comparison of the performances in terms of error and response times of the neural networks trained by various training functions. According to their results, the best performing training function is Levenberg-Marquardt back-propagation algorithm when the neural network has one hidden layer of 8 hidden units. The authors also show that the neural network performs similar to NNSS algorithm; but, it uses less memory, and it has better responsiveness, i.e., its response time is shorter than the response time of NNSS algorithms.

2.3.1.4 Probabilistic Calibration-Based Approaches

Joint Clustering (JC) Technique is a probabilistic approach which is based on the joint distribution of RSS values collected from multiple RPs [55]. In the calibration phase, given an indoor region covered by multiple RPs (access points with known locations), the system collects RP RSSs at various locations and constructs a histogram-based radio map.

For a given training location l, the probability that RSS reading at RP_i is s_i equals to the histogram of RP_i at s_i,

$$P(RP_i = s_i|l) = Hist_{l,RP_i}(s_i). \tag{2.3}$$

Then, the histograms are used to estimate the joint distribution of the RSS values received from k RPs at each training location l. When RPs are assumed independent, the problem of estimating the joint probability distribution becomes the problem of estimating the marginal probability distributions as:

$$P(RP_1 = s_1, RP_2 = s_2, \ldots, RP_k = s_k) = P(RP_1 = s_1)P(RP_2 = s_2)\ldots P(RP_k = s_k) \tag{2.4}$$

Therefore, at the calibration phase, the radio map stores the joint probability distribution of RSSs received from k RPs at each training location. It is important to choose k such that all locations are covered by at least k RPs most of the time.

After calibration, location clustering is performed to reduce the size of the search space and the computational requirements of the algorithm. In location clustering the locations that share q RPs with the highest RSS are grouped together by putting them in the same cluster.

In the location estimation phase, RSS samples from RPs at an arbitrary location are taken. Then, the joint distributions are used to find the most probable location given the run-time RSS values. q RPs with the largest RSS values are used to determine one cluster to search within for the most probable location. The locations in this cluster are called target locations. Then, the location estimation algorithm looks for location l in the target locations, which maximizes the probability that being at location l given the run-time RSS samples, *samples*, from the RPs, $P(l|samples)$. Using *Bayes' theorem*,

$$argmax_l[P(l|samples)] = argmax_l\left[\frac{P(samples|l)P(l)}{P(samples)}\right] \quad (2.5)$$

$$= argmax_l[P(samples|l)P(l)], \quad (2.6)$$

assuming $P(samples)$ is independent of l.

If the user's profile is not known, then all locations for a user is equally likely, then $P(l)$ can be factored out from the maximization process. And the remaining term can be calculated by using:

$$argmax_l[P(l|samples)] = argmax_l[P(samples|l)],$$

$$P(samples|l) = \prod_{n=1}^{|samples|} \prod_{i}^{k} Hist_{l,RP_i}(samples(n)),$$

where k is the number of RPs to be included in the joint distribution. k RPs with the largest average RSS are chosen among a total of N RPs.

Therefore, in [55], Bayes' theorem is used to estimate the probability of each location within the cluster given the run-time samples and the radio map built during the offline phase. The most probable location is reported as the estimated user location l.

The radio map of the ***Incremental Triangulation (IT) Technique*** technique is the same as the radio map of JC technique except location clustering [55]. JC technique calculates the probability of a location using k RPs all at the same time, using k operations per sample. IT uses the RPs incrementally, one after the other, until it can estimate the location with certain accuracy, using a predetermined threshold.

Given the run-time observations from each RP, IT calculates the probability of each location in the radio map set given the sample vector from the RP with the strongest average RSS alone. If the probability of the most probable location ($argmax_l[P(l|samples)]$) is "significantly" higher than the probability of the second most probable location according to a measure defined in the algorithm, the most probable location is returned as the location estimate, after consulting only one RP. If this is not the case, for the RP with the second strongest average RSS, the same process is repeated, but only for the set of candidate locations obtained from the first RP. This process of calculating the probabilities and determining the significance of the most probable location is repeated incrementally, for each RP in order, until the location can be estimated or all RPs are consulted. In the latter case, the algorithm returns the most probable location in the candidate list that remains after consulting all the RPs.

This iterative process leads to a multi-level clustering which reduces the search space significantly at each iteration; and hence, leads to less computation than the JC approach. However, treating each RP incrementally, instead of using the joint distribution, causes the loss of some information in IT and a lower accuracy than the JC technique.

In another probabilistic approach Youssef proposes ***Horus*** system which exploits the *samples correlation* in calculating the parameters of the distribution of the average of n samples for each RP in the radio map [42]. This algorithm assumes RSS has *Gaussian* distribution. Since the individual distribution of each sample follows Gaussian distribution, the probability distribution of the average of n samples also follows a Gaussian distribution. This algorithm exploits samples correlation in order to estimate the variance and the mean of this Gaussian distribution. The mean of the distribution of the average of n samples, is the same as the mean of the distribution of each sample, $E(A(n)) = E(s)$. The variance of the averaging process, $Var(A(n))$ depends on the variance of the original process, $Var(s)$, the number

of samples, n, and the independence coefficient, α. $Var(A(n))$ is always less than or equal to the variance of the original process, $Var(s)$, being equal in case α equals one or if the RSS samples are independent. Intuitively, the lower the variance of the RSS distribution at each location, the better the ability to discriminate between different locations and the better the accuracy. Therefore, this approach takes the independence of samples into account. Youssef [42] showed that as the number of samples increases, the estimation of the distribution of the average of n samples becomes worse due to the wrong independence assumption. Before the location estimation phase, training locations are grouped into clusters according to the RPs that cover the locations. The objective is to reduce the computational requirements of the system.

In location estimation phase, the algorithm obtains the average of n run-time samples from an RP, and calculates the probability of each radio map location given this value of the average using the distribution of the average of n samples calculated during the offline phase. Then, the location estimation phase calculates the probabilities for each RP, and then, returns the radio map location that has the maximum probability as the estimated location.

In ***Bayesian Inference (BI) algorithm***, instead of using only the histograms of RSS measurements, both frequency count from RPs and RSS histograms are employed [56]. In the calibration phase, a sensor map of the environment is built by gathering data at various predefined points of the environment. The state space S consists of the points defined by a tuple $s_i = (x, y, \theta)$ coordinates and the orientation. Each observation o consists of the frequency count and RSS measurements from all base stations (BS). At each state s_i, $i = 1, \ldots, n$, an observation is taken, and two histograms are built for each BS for that point: the distribution of frequency counts over the sampled observations, and the distribution of observed RSSs. Based on these histograms, two conditional probabilities are calculated: the probability that the frequency count for the j^{th} BS is a at state s_i, $Pr(f_j = a|s_i)$; and the probability that BS b_j has RSS RSS_j at state s_i, $Pr(RSS_j|b_j, s_i)$. By multiplying these conditional probabilities the conditional probability of receiving a particular observation,

$$o =< k, f_1, \ldots, f_N, (b_1, RSS_1), \ldots, (b_k, RSS_k) >,$$

at a state s_i is obtained, when there are N BSs and k readings from BSs. This probability is,

$$Pr(o|s_i) = \left(\prod_{j=1}^{N} Pr(f_j|s_i)\right) \cdot \left(\prod_{j=1}^{k} Pr(RSS_j|b_j, s_i)\right). \tag{2.7}$$

In location estimation phase, the prior state distribution, π, is assumed uniform. After making the observation, Bayes' rule is used to compute the probability distribution over the states, π'.

$$\pi'_i = \frac{\pi_i Pr(o_j|s_i)}{\sum_{i=1}^{n} \pi_i Pr(o_j|s_i)} \tag{2.8}$$

Then an appropriate candidate location is chosen as the estimated location. When the location with the highest probability is chosen, the observed error is less than 1.5 meters with probability 0.77.

In order to improve the results, a sensor fusion technique is used after obtaining π'. In sensor fusion, the probability distributions found by the Bayes' rule are post-processed. A simplistic sensor fusion technique transforms the probability distribution over the states to $\pi'_i = (\pi_i + u_1)(\nu_i + u_2)$, where π is the prior distribution, ν is the probability distribution computed with the Bayes' rule, and u_1 and u_2 are small constants representing artificial uniform distributions. This technique improved results by 8%.

2.3.2 Calibration-Free Localization

Laborious calibration efforts in calibration-based localization systems has led researchers to develop calibration-free localization algorithms.

A calibration free system does not require an offline phase where the RSS vectors are mapped to training locations. Learning is realized during the online phase with or without the location information requirement for the training points. The majority of calibration-free algorithms takes as input the online measurements of RSSs between APs, and between a mobile device and its neighboring APs. RSS measurements among APs are used to capture the real time behavior of RSS and to create a mapping between the RSS measure and the actual geographical distance to anchors. For online training, for most of the calibration-free solutions no human intervention or client side effort is needed, and learning can be

repeated continuously. Therefore, calibration-free systems eliminate laborious offline RSS measurements, while they are robust against the changes in the environment due to their adaptation capabilities. In this section state-of-the-art calibration-free localization systems are discussed in two groups, deterministic and probabilistic.

2.3.2.1 Deterministic Calibration-free Approaches

Triangular Interpolation and eXtrapolation (TIX) is based on inter-AP distance mapping curves which are generated online to find the approximate distances between the mobile and RPs [1].

This localization system uses the mobile RSS measurements which are the RSSs measured at the mobile device from all APs, and also the inter-AP measurements which are the RSSs measured at all APs from all the other APs. Inter-AP measurements are used to generate multiple distance mapping curves (RSS vs. distance curve) for each AP observed. Generating equation of a line requires at least two points. The first point is obtained by using inter-AP RSS measurement and each AP's location, which is known to the system. The second point is the maximum RSS value measured at the minimum measurable range, which is also known to the system. When there are four APs collecting RSSIs from all other APs excluding itself, the system generates three mapping curves for each AP. Figure 2.5 shows an example of distance mapping curves for each AP observed at AP_2, when there are four APs in total.

The mobile RSS measurements and the appropriate mapping curves from each AP are used to find the approximate distance between the mobile and the APs. In order to select the appropriate curves, a simple heuristic called *Proximity in Signal Space (PSS)* is used. According to this, the curves that are generated with the highest RSS are chosen for each AP. For the example shown in Figure 2.5, the AP with the highest RSS is AP_2, $S_2 > S_1 > S_3 > S_4$, where S_i is the RSS measured at the mobile from AP_i. Using the curves and the mobile RSS measurements, the approximated distances between the mobile and each AP is found as d_1, d_3 and d_4. d_2 is found from the curves generated at AP_1, since it has the second highest RSS. The approximated distances between the mobile and APs and location coordinates of APs are input to TIX algorithm. As in trilateration, in order to give a 2-dimensional result, TIX needs at least three APs. When there are more than three, TIX chooses the best three APs, highest three mobile-measured RSSIs.

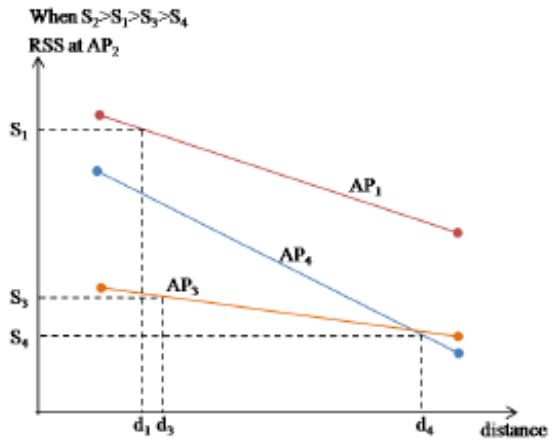

Figure 2.5: RSS distance curves for AP_2 when there are four APs in total

TIX first forms a triangle, whose vertices are the locations of APs, then, uses the internal or external dividers of the sides of this triangle to determine the final location estimate. In the triangular interpolation, the sides of the triangle are internally divided by points D_1, D_2, and D_3. The centroid of triangle $D_1D_2D_3$ gives the final location estimate. In triangular extrapolation, two sides of the triangle are externally divided by points D_1 and D_3 while the third side is internally divided by D_2. Again, the centroid of triangle $D_1D_2D_3$ gives the final location estimate. Triangular interpolation and extrapolation are shown in Figures 2.6 and **??**.

Similar to TIX algorithm, ***Truncated Singular Value Decomposition (SVD) Technique*** takes as input the online measurements of RSSs between APs, and between a mobile device and its neighboring APs [26]. RSS measurements among APs are used to capture the real time behavior of RSS and to create a mapping between the RSS measure and the actual geographical distance to anchors. The mapping is created online with the use of *truncated singular value decomposition (SVD)* techniques. The objective is to mitigate the effect of measurement errors and to characterize the anisotropic relationship between RSSs and geographic distances, while retaining as much environmental information as possible. The location of a wireless mobile device can be calculated with the signal-distance map (SDM) and the user-measured RSSs between itself and its neighboring APs via trilateration.

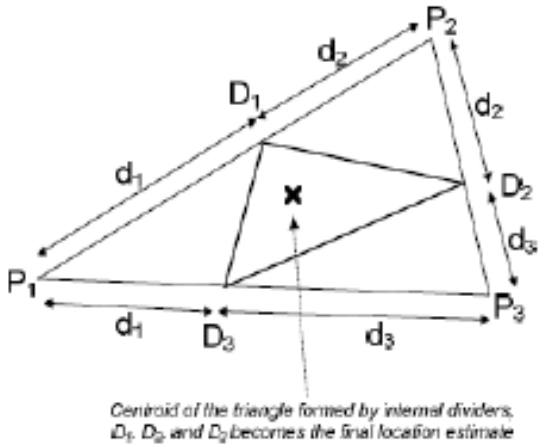

Figure 2.6: Triangular interpolation [1]

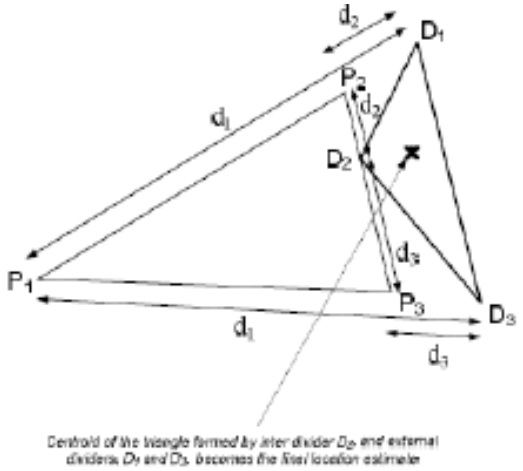

Figure 2.7: Triangular extrapolation [1]

In this work, localization is considered as an embedding problem that maps the set of objects into an embedding space. Based on this concept, each mobile device to be localized has two coordinates in Lipschitz embedding spaces that correspond to the RSS measure and the Euclidean distance between itself and neighboring APs, respectively. According to this, the RSSs measured from a node to APs define the coordinate of the node. Given that there

exist m APs, the coordinate of a node i in an m-dimensional Lipschitz embedding space is represented by a signal vector, $s_i = [s_{i1}, s_{i2}, \ldots, s_{im}]^T$, where s_{ij} is the RSS emitted by the jth AP and received by node i. Then, RSSs measured between APs can be represented by an m-by-m RSS matrix $\mathbf{S} = [s_1, \ldots, s_m]$, whose ith column is the coordinate of the ith beacon node. Similarly, they define the geographic distance vector and matrix as $d_i = [d_{i1}, d_{i2}, \ldots, d_{im}]^T$ and $\mathbf{D} = [d_1, \ldots, d_m]$.

As a good candidate for the SDM, an optimal linear transformation matrix $\mathbf{T}$ is employed. In $\mathbf{T}$ the geographic distance from a mobile device node to an AP is represented as a weighted sum of the RSSs to all the APs. The element t_{ij} of matrix $\mathbf{T}$ represents the effect of RSS measurement to the jth AP on the geographic distance to the ith AP, and can be considered as a scaling factor from RSS to distance. Each row vector t_i can be obtained by minimizing the following error function:

$$e_i = \sum_{k=1}^{m} (log(d_{ik}) - t_i s_k)^2 = ||log(d^{T_i}) - t_i S||^2.$$

Then the least square solution obtained for t_i and SDM matrix $\mathbf{T}$ are $t_i = log(d_i{}^T)\mathbf{S}^T(\mathbf{S}\mathbf{S}^T)^{-1}$ and $\mathbf{T} = log(\mathbf{D})\mathbf{S}^T(\mathbf{S}\mathbf{S}^T)^{-1}$, respectively. SDM retains the effects of physical wireless characteristics on the RSS to all the APs, and hence well characterizes the anisotropic relationship between RSSs and geographic distances.

Having RSS measurements between itself and its neighboring APs, a mobile device node n can calculate the distance from itself to APs by matrix multiplication,

$$d_n = exp(T s_n).$$

Then, a lateration algorithm can be used to estimate the location. In this paper, the *simple descent gradient method* which minimizes the error between the actual location and the estimated location is used. This method minimizes the following equation:

$$\epsilon_i = \frac{1}{2} \sum_{i=1}^{m} (f_d(\tilde{x}, x_i) - \tilde{d_i})^2,$$

where $\tilde{x}$ and $\tilde{d_i}$ are the estimated location of the mobile node and the estimated distance to the ith AP computed by SDM, respectively. $f_d(\tilde{x}, x)$ denotes the Euclidean distance

between locations x and $\tilde{x}$. By differentiating the objective function, an iterative equation for updating $\tilde{x}$ is found as:

$$\tilde{x}[k+1] = \tilde{x}[k] + \alpha \sum_{i=1}^{m} (1 - \frac{\tilde{d}_i}{f_d(\tilde{x}, x_i)})(\tilde{x}[k] - x_i),$$

where α is a constant step size.

Truncated SVD pseudo-inverse method is used to reduce the measurement noises in matrix $\mathbf{T}$. According to SVD, $\mathbf{S}$ can be expressed in a matrix form as a function of Σ, $\mathbf{U}$ and $\mathbf{V}$, where Σ is a diagonal matrix, $\Sigma = diag(\rho_1, \ldots, \rho_w)$, and $\mathbf{U} = [u_1, \ldots, u_m]$ and $\mathbf{V} = [v_1, \ldots, v_m]$ are column and row orthogonal matrices, and w is the rank of matrix $\mathbf{S}$. ρ_s are the singular values of $\mathbf{S}$. Then the pseudo-inverse of $\mathbf{S}$ can be written as $\mathbf{S}^+ = \mathbf{S}^T(\mathbf{S}\mathbf{S}^T)^{-1} = \sum_{i=1}^{w} \frac{1}{\rho_i} v_i u_i^T$. Furthermore, the truncated pseudo-inverse of $\mathbf{S}$, $\mathbf{S}_\gamma^+ = \sum_{i=1}^{\gamma} \frac{1}{\rho_i} v_i u_i^T$. Then, SDM can be expressed as $\mathbf{T} = log(\mathbf{D})\mathbf{S}_\gamma^+$, and can be updated online. Then, the geographic location of the mobile n can be determined by lateration algorithms as explained above.

The only calibration in the deployment phase is the RSS measured between two co-located APs. In this work authors measure the RSS between two co-located APs for 5 minutes and take the median value of the measurements as the self-RSS. In the implementation, in order to effectively deal with short-term variations in RSS, SDM uses a median filter that takes the median value of RSS samples measured for a time duration T_s, which is set to 60 seconds in the experiments.

REDPIN system [57] proposes a ***calibration free fingerprinting*** approach that does not require an offline calibration phase but incorporates the training into the usage of the system, making the system adaptive against environment changes. As a difference from the other discussed algorithms, the objective of REDPIN is to identify the room the mobile is located in rather than identifying the coordinates of the mobile. Therefore, this is a coarse-grained localization system.

At the initialization, the system does not know anything about the building, or the locations of APs, etc. REDPIN allows the users of the system create and manage the locations in a collaborative way. In order to identify the room, the application on the mobile device measures the RSS of the currently active GSM cell, the RSS of all WiFi APs as well as the Bluetooth identifier of all non-portable Bluetooth devices in range. Then, these

measurements are sent to a central server which will compute the Euclidean distance between the current measurements and the fingerprints at the database. The system will return the best match. If the device cannot be located, the user will be prompted to name the place of his or her current location and to indicate the appropriate position on the floor plan. Therefore, the time it takes to get at least one fingerprint for every room depends only on how active users are in contributing to the system and on their mobility. This approach works if there are enough users in the system contributing to the training of the system.

2.3.2.2 Probabilistic Calibration-Free Approaches

In [58], authors proposes a probabilistic calibration free approach that is based on ***wireless sniffers*** and reference APs, and construction of a radio propagation map (RPM) containing the grid positions in a given space associated to the probability that receiving a RSS tuple from k sniffers given the position. During the RPM construction, the sniffers detect wireless mobile devices and record RSS values from them. The sniffers capture frames for each detected transmitting device during the capturing interval (CI), which is set to 1 second in the experiments, taking advantage of high auto-correlation between consecutive RSS values in this interval [58]. The average value $\overline{RSS}$ of all RSSIs measured in the CI is computed and sent to the database. In addition, sniffers (i) capture M beacon frames from each reference AP (n) and calculate the mean and the standard deviation of M RSSIs; and then send the pair $(\mu_{i;n}, \rho_{i;n})$ to the database. The sniffers' and APs' locations are known to the system. Therefore, when there are k sniffers and n APs, n times k pairs are sent to the database.

The position l which maximizes $P(l|s)$, the probability that the wireless device is at location l given the RSS vector $\mathbf{s} = (s_1, \ldots, s_k)$ measured by k sniffers, is the estimated location. The location server, having the locations of reference APs and the RSS readings from sniffers, associates each grid position, $l = (x, y) \in L$ with $P(\mathbf{s}|l)$, the probability that having vector $\mathbf{s}$ given the location is l. Having $(\mu_{i;n}, \rho_{i;n})$, a Gaussian distribution is used to calculate $P(\mathbf{s}|l)$ as follows:

$$P(\mathbf{s}|l) = \frac{1}{\rho_{(l)}\sqrt{2\pi}} exp\left(-\frac{(\mathbf{s} - \mu_{(l)})^2}{2\rho_{(l)}^2}\right),$$

where $\rho_{(l)}$ and $\mu_{(l)}$ are the standard deviation and the expected RSS value given the object is located at l. $\mu_{(l)}$ is estimated by using a large scale propagation model and the distances between sniffers and APs, and sniffers and transmitter located at l. $\rho_{(l)}$ is assumed equal to $\rho_{i;n}$.

Since this model uses reference APs and sniffers to model the RSS behavior, it is adaptive to the changes in the environment, and the radio propagation map can be updated when a deviation in the mean of RSS measured at sniffers occurs.

Then the probability that a mobile device is at a location given the RSS vector $\mathbf{s}$ from the sniffers is calculated using Bayes rule as follows:

$$P(l|\mathbf{s}) = \frac{P(\mathbf{s}|l)P(l)}{\sum_{l' \in L} P(\mathbf{s}|l')P(l')},$$

where l' is all locations in the RPM grid. Also, the measurements at each sniffer are assumed independent, and the user profile $P(l)$ is assumed uniformly distributed. Therefore,

$$P(l|\mathbf{s}) = P(l|s_1, \ldots, s_k) = \prod_{i=1}^{k} P(l|s_i).$$

Then, the most probable location, $l = argmax[P(l|s)]$, is returned as the estimated location of the wireless device.

Against the possibility that several different grid points can present similar values and cause estimation errors, the estimation window technique is used. When the estimation window is 10, the estimation is done 10 times, and the estimation with the highest probability is returned as the estimated location. Despite the increased accuracy of this technique, response time increases with the increasing estimation window. Another technique to increase the accuracy without increasing the response time is to calculate the weighted average of the locations. According to this technique, the higher the probability the more weight a location gets.

All localization algorithms, we previously cited require location information of either wireless sniffers or reference points or training locations in calibration phase. Madigan et. al. used *Bayesian Graphical (BG) models* for location estimation and showed that a hierarchical Bayesian approach can provide accurate location estimates without any location information in the training data [2].

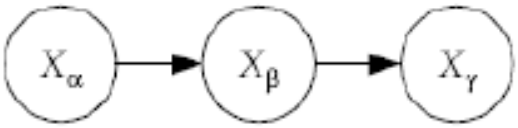

Figure 2.8: A simple acyclic directed graph model [2]

In a BG Model, model parameters are random variables and appear as vertices in the graph, and edges represent the conditional independence relationships between the random variables. Figure 2.8 shows an example of a directed graph, which represents the assumption that X_γ and X_α are conditionally independent given X_β. The joint density of the three variables then,

$$P(X_\alpha, X_\beta, X_\gamma) = P(X_\alpha)P(X_\beta|X_\alpha)P(X_\gamma|X_\beta).$$

A more general expression is for a joint density of random variables X_v, $v \in V$ is given by,

$$P(V) = \prod_{v \in V} P(v|pa(v)),$$

where $pa(v)$ represents the parents of a vertex v.

When some variables are discrete and others are continuous in a graphical model, a closed-form Bayesian analysis is not available. In this case, *Markov chain Monte Carlo (MCMC)* simulation method is employed for the analysis. Gibbs sampler and MCMC algorithm for BG models are explained in the paper.

The Gibbs sampler starts with some initial values for each unknown parameter in the graph, and then cycles through the graph by simulating each variable v in turn from its conditional probability distribution, given all the other quantities, denoted as $V \backslash v$. Then, the simulated v replaces the old value, and simulation shifts to other quantity. After sufficient iterations, MC reaches its stationary distribution, and then the future simulated values for

the random values are monitored. The analytic summaries of these monitored values are used to estimate the inferences concerning the unknown quantities.

The full conditional distribution for any vertex v depends on a prior terms and a set of likelihood terms, and is equal to

$$P(v|V\backslash v) = p(v|pa(v)) \prod_{w \in child(v)} p(w|pa(w)).$$

Thus, when sampling from the full conditional for v, only vertices which are parents, children or parents of children of v is required for local computations.

In the paper, two BG Models are proposed and the results for them are shown. The first model is a ***Non-Hierarchical Bayesian Graphical (NHBG) Model*** which requires location information along with the RSS information from the training locations. In Figure 2.9, the first graphical model is shown. In this model, X and Y represent location, and they are assumed uniformly distributed, $X \sim Uniform(0, L)$ and $Y \sim Uniform(0, B)$, where L and B are the length and the breadth of the building respectively. Vertex D_i represent the Euclidean distance between the location specified with X and Y and the ith AP, and it is deterministic. The vertex S_i represents the RSS measured at (X, Y) with respect to ith AP, and it has a Gaussian distribution, $S_i \sim N(b_{i0} + b_{i1} log(D_i), \tau_i)$. The mean and the standard deviation of this distribution depend on other Gaussian distributions, $b_i 0 \sim N(0, 0.001)$ and $b_i 1 \sim N(0, 0.001)$ and a precision value (τ).

The second model, ***Hierarchical Bayesian Graphical (HBG) Model***, only requires RSS measurements. This model incorporates the knowledge that RSS decays approximately linearly with the logarithm of distance with the knowledge that the different APs behave similarly, because each AP has similar physical processes. With this model, the training data comprise vectors of RSSs with unknown locations; the location coordinates, X and Y, in the model become latent variables. Figure 2.10 shows this graphical model. The difference from the previous model is that, the parameters $b_{i0} \sim N(b_0, \tau_{b0})$ and $b_{i1} \sim N(b_1, \tau_{b1})$ depend on other Gaussian and Gamma distributed parameters, $b_0 \sim N(0, 0.001)$, $b_1 \sim N(0, 0.001)$ and $\tau_{b0} \sim Gamma(0.001, 0.001)$ and $\tau_{b1} \sim Gamma(0.001, 0.001)$.

After training the model, MCMC algorithm estimates the parameters and produces location estimates. According to the results, the second model without the location information,

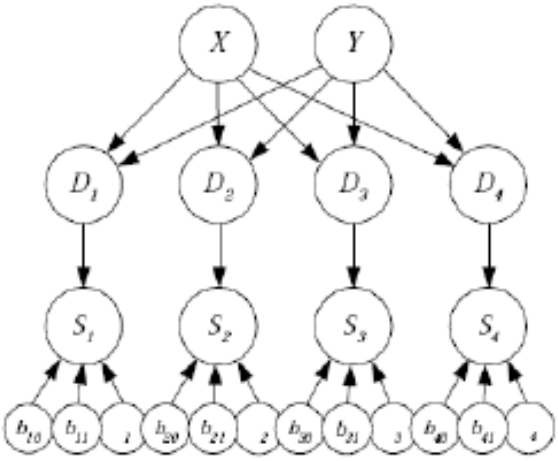

Figure 2.9: Nonhierarchical Bayesian Graphical Model [2]

from about 10 training vectors onwards, performs almost as well as the LEASE technique [46] which has trained on data with complete location information for each RSS vector.

This model, despite not eliminating the training phase all together, drops the location data requirement, and; therefore, provides significant practical benefits: gathering RSSs vectors without the corresponding locations does not require human intervention; suitably instrumented APs or sniffing devices can solicit RSS measurements from existing Wi-Fi devices and can do this repeatedly at essentially no cost.

2.3.3 Comparison of Calibration-based and Calibration-free Localization

Table 2.16 presents a summary for the deployment and training features and the results in terms of the accuracy and responsiveness of the systems described in the previous sections. According to this table, it is not trivial to compare different localization systems due to the differences in the deployment and differences between the conditions that the systems trained and tested. For example, the size and the shape of the area the system tested, the number of APs used, the number training locations used in calibration can change from system to system. In this section, four different error metrics are used to compare different

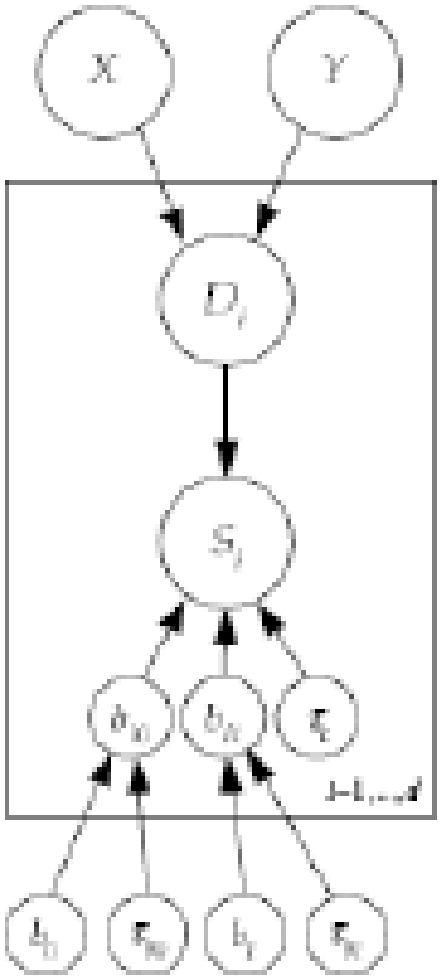

Figure 2.10: Hierarchical Bayesian Graphical Model [2]

localizations systems as fair as possible. For the sake of simplicity, shorter names are given to the localization algorithms, they can be found in Table 2.16.

The first metric considers the median localization error, m, and the size of the experiment area, A. The objective of this metric is to reveal the relationship between the localization accuracy that can be achieved in a given localization service area.

$$e_1 = \frac{m}{A}. \tag{2.9}$$

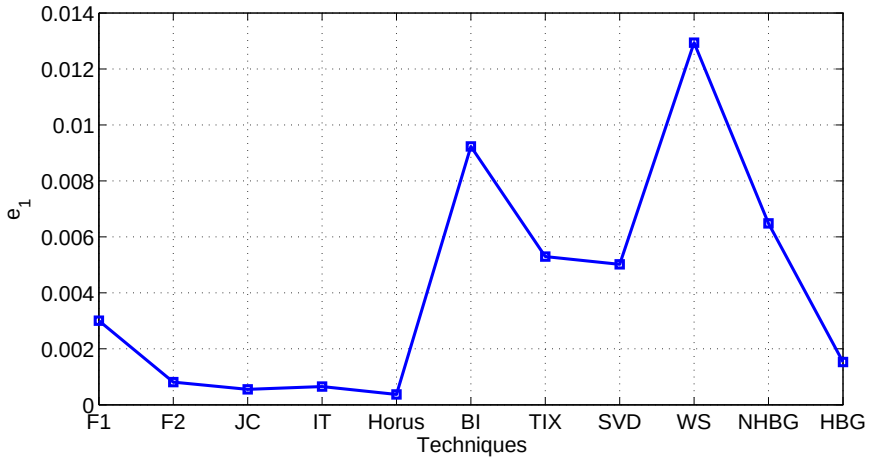

Figure 2.11: Median error per experiment area for calibration-based and calibration-free techniques

Figure 2.11 shows the performance of various localization algorithms according to this error metric, e_1. This figure implies that when compared to calibration-free techniques, calibration based techniques (except NN) results in lower error on the average per experiment area. However, this metric does not take the cost due to calibration efforts into account.

The second metric considers the number of training locations, t, in addition to m and A [46],

$$e_2 = \frac{t^i m^j}{A}. \tag{2.10}$$

The parameters i and j are used to tune the error metric to emphasize the cost of calibration, t, or the raw error estimation, m, respectively.

We first use this metric to show the effect of calibration on localization accuracy rather than comparing different systems. Figure 2.12 is plotted using the results of the RADAR [3] system. According to this figure, as the number of training locations decreases, error metric increases. However, note that when the emphasis on the median error is not high, the performance does not suffer greatly from decreasing the number of training locations until it reaches a point where the number of training locations is very low. Another observation

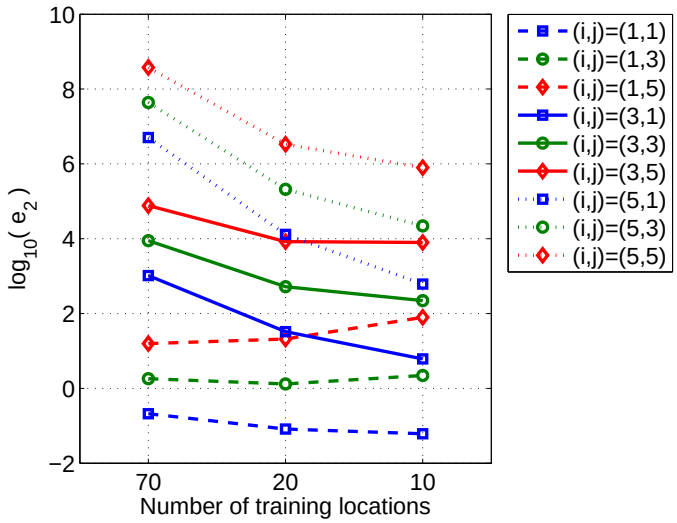

Figure 2.12: Error metric vs. the number of training locations in RADAR [3]

is that as the emphasis on the median error (j) is increasing, the error metric increases with a higher slope with decreasing t. The effect of the number of readings per location is investigated in [47]. According to this work, the accuracy becomes worse although it does not suffer significantly even when the time spent in each location is only ten seconds when compared to 60 seconds. Therefore, the performance of a calibration-based method depends on how much effort is spent at the calibration phase, although parameters can be reduced without a correspondingly significant reduction in the performance.

Figure 2.13 shows the effect of the emphasis on the median error, j, on the error metric for different calibration-based systems. According to this figure, when $j = 1, 2$ the best algorithm is GD40, when $j = 3, 4$, the best algorithm is BI52, and the probabilistic approaches outperform deterministic approaches.

The second metric is useful in comparing different calibration-based systems, but it cannot be used to compare calibration-based and calibration-free systems. Therefore, the third metric is considered to include the number of infrastructure used, a, such as the number

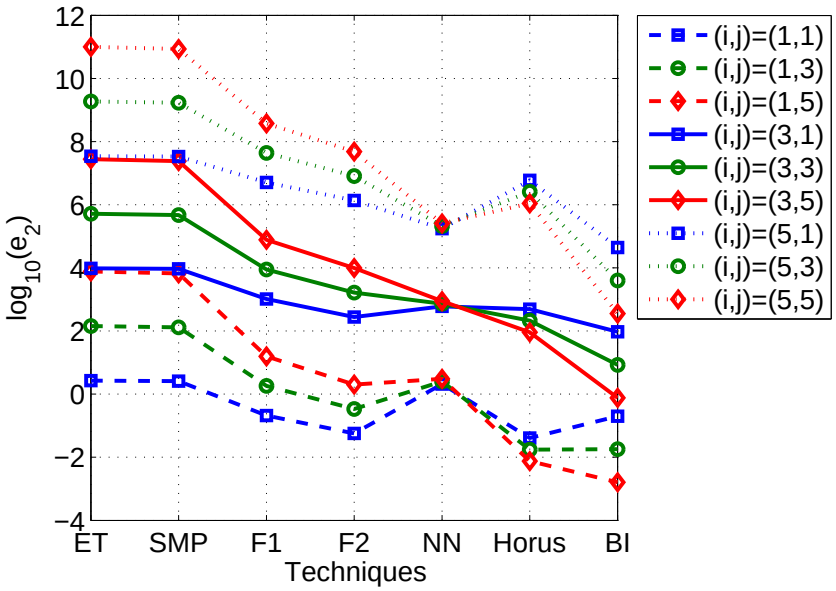

Figure 2.13: Logarithm of error metric for calibration-based techniques

of APs, instead of the number of training locations,

$$e_3 = \frac{m^j a^k}{A}, \tag{2.11}$$

where k is used to tune the emphasis of the deployment cost due to hardware.

Figure 2.14 shows the effect of the emphasis on the error median in calibration-based and calibration-free techniques. When $j = 1$, the best algorithm is F24, $j = 2$, the best algorithm is GD40, and otherwise, the best algorithm is BI52. All of these algorithms are calibration-based algorithms. Among the calibration-free algorithms, HBG25 is the best performing algorithm.

And, Figure 2.15 shows the effect of the number of infrastructure used in the systems. According to this, when $j = k = 2$, calibration-based probabilistic algorithms, GD40 and BI52 perform better than other algorithms, and otherwise, BI52 outperforms others. HBG25 and WS53 perform better than the other calibration-free approaches, and their performances are comparable to calibration-based approaches.

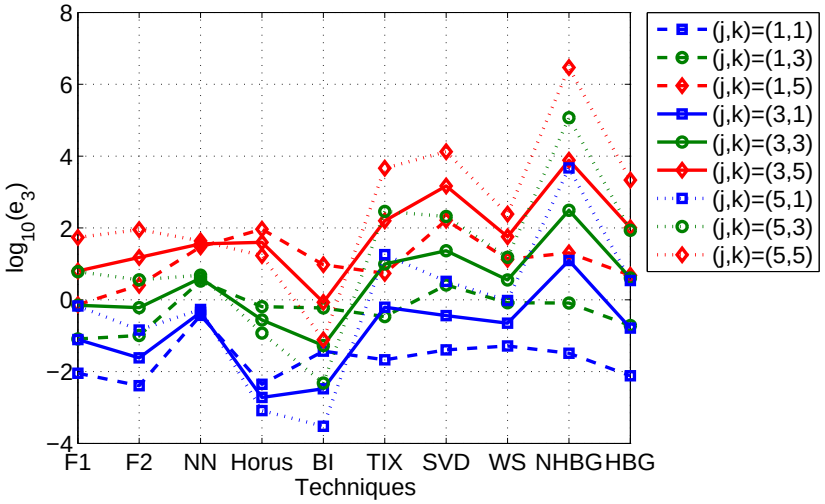

Figure 2.14: Logarithm of error metric for calibration-based and calibration-free techniques, for changing values of j

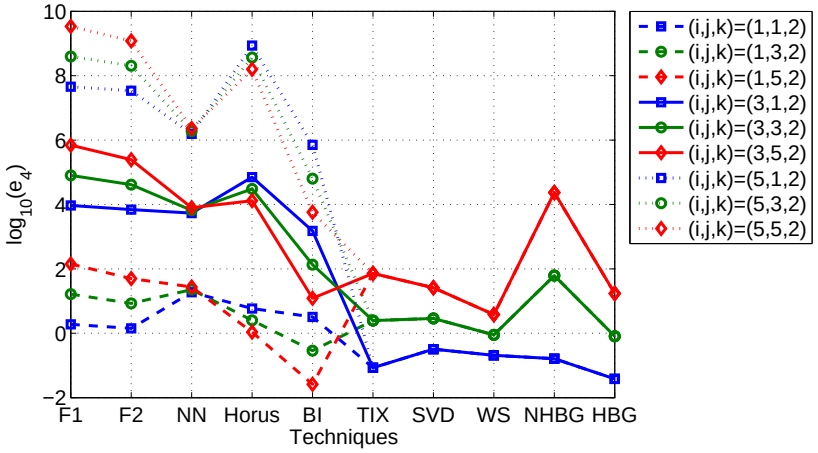

Figure 2.15: Logarithm of error metric for calibration-based and calibration-free techniques, for changing values of k

In order to comment on the scalability of localization techniques, a forth metric,

$$e_4 = \frac{t^i m^j a^k}{A}, \tag{2.12}$$

is used; and for calibration-free techniques the number of training locations is assumed 1. According to Figure 2.15, calibration free systems are more desirable than calibration based systems in terms of the scalability. The reason is the difficulty of calibrating a whole system in a very large area containing high number of infrastructure. These systems involve a steep upfront cost and effort to deployment, and add significantly to the complexity of maintaining the model. The installation cost increases due to the wages of people who work in the measurement collection for the calibration phase. In addition, it has been shown that a reasonable calibration result can be obtained by collecting training data roughly uniformly from each location in the whole area [3]. This procedure consists of physically moving a wireless device over each training point and capturing RSS values from APs, and it can be considered as a significant barrier to wider adoption of such location methods. In previous studies the authors reported a 4-9 hours of calibration. Moreover, the environment may not remain consistent due to the addition or replacement of APs and structural changes that would affect radio propagation. In addition to the structural changes, the channel values can change significantly due to fading or the changes in the environmental conditions, and cannot be assumed constant from calibration to estimation period [30]. In most of the calibration-based studies, a bunch of measurements are taken and divided into two, in order to use the first part in calibration and the second part in estimation. This implies test results cannot be relied on as a simulation of real-life example.

A more reliable calibration may be reached by periodic calibrations or by an extensive data collection held in different times of day and different days of week. Therefore, calibration repeatedly consumes human labors and creates significant maintenance and scalability issues. On the other hand, calibration free techniques can provide more robust and scalable solutions due to their adaptation ability against the changing conditions of wireless environment. This ability stems from the online learning process of calibration-free techniques. This process is more cost effective, since it does not require human intervention or any kind of labor.

Calibration-based and calibration free techniques cannot be compared with respect to responsiveness due to unclear experimental information. However, intuitively, since the learning in calibration-free techniques is held in real time, the number of samples used in learning affects the latency of calculating the location estimate.

All of the works cited in this chapter are interested in location estimation techniques aiming better accuracy; however, they lack the experiments and discussion on the effect of the background traffic on the accuracy of the localization system. [59] implements a simple RSSI-based localization system on a 802.15.4-based sensor network testbed, investigates the effect of background IEEE 802.11 traffic on localization error. According to the results, the 80th-percentile of the localization error may increase by 141% at worst in an office building with active use of WiFi for data due to losses in RSSI readings as the beacon messages collide with background traffic. Since we cannot find this kind of analysis in other studies, we are not able to compare the calibration-based and calibration-free systems with respect to their reliability under background traffic creating interference.

2.4 REFERENCE NODE PLACEMENT STRATEGIES

Irrespective of the localization algorithm used, the reference node placement strategy adopted in a localization system significantly affects the accuracy, practicality and the cost of the system directly. In this section, we will briefly review the literature on reference node deployment.

In [60], an empirically adaptive beacon placement scheme for a proximity based localization algorithm has been proposed. The objective is to adjust the number and locations of beacons (reference points) to adapt to the noisy and unpredictable environmental conditions. This is a calibration based approach as it requires exploration of the localization area conditions by a mobile human or robot agent. According simulations results, beacon density can significantly effect the mean localization error; however, increasing the density more than 0.01 per m^2 does not help improve accuracy. The effect of relationship between the density and beacon transmission range was discussed.

Reference node placement strategies presented in [61], [62] and [63] assume a fixed transmission range. [61] is proposing landmark placement strategies in topology based localization in sensor networks. In this localization scheme, the coordinates of sensor nodes are given depending on the hop count distance to multiple landmarks. In [61] and [62] it is shown

that placing equally spaced RNs on the perimeter of the network yields better accuracy. In [63], a RN placement strategy for a triangulation based localization algorithm is proposed. In this strategy, the RNs must be placed so that they will form a equilateral triangle, where one side of the triangle is equal to the transmission range.

On the other hand, [64] uses RF profiling with an offline calibration, and argues that the effect of the location and the number of calibration locations are more important than the effect of the number and the locations of RNs. In [65], a range based landmark placement optimization algorithm is proposed. A pre-specified deployment shape is stretched/shrinked iteratively to minimize the localization error. According to this work, for small number of landmarks, simple shapes such as equilateral triangles and squares result in better localization performance. In [66], Zheng et.al. proposed a strategy for placing RNs to improve the accuracy of ToA based localization.

An interesting research area is related to connectivity and coverage problems in wireless ad hoc networks. In [67], authors discuss the importance of exposure of a mobile object to the sensor network over a given period of time. Depending on the exposure requirement and with the heuristics to calculate the minimal exposure path representing the worst case exposure-based coverage, the parameters of uniform distribution for sensor placement are given. [68] discusses the sensor deployment problem when sensor locations are unknown and random due to airdropping. The objective is to optimize the number and coverage of sensors. [69] discusses the incremental deployment algorithm for maximum coverage for mobile sensor networks. The location of each sensor added to the network is calculated by using the information from sensors deployed previously. This work differs from [60] because network nodes are assumed to be equipped with sensors that require line-of-sight to operate such as cameras.

2.5 MULTIPLE FLOOR LOCALIZATION ALGORITHMS

Most localization schemes proposed in the literature have predominantly focused on single-floor or 2-D localization. In indoor areas with multiple floors, it is important to distinguish

between floors and also localize mobile nodes within a floor accurately.

Otsason et al. used RSS fingerprints of GSM signals [70] in three multi-floor buildings and showed that they can achieve 2.48m - 5.44m within-floor accuracy, and correct floor classifications between 89% and 97% of the time. Varshavsky et al. also proposed a GSM-fingerprint based localization system called SkyLoc for floor identification in tall multi-floor buildings [71]. Using machine learning techniques, Skyloc could determine the correct floor in up to 73% of the cases and was within 2 floors of the correct floor in 97% of the cases. Rehman et al. presented CILoS, a fingerprinting localization system based on the signal delay [72] of multiple CDMA channels. This system could achieve a median accuracy between 4.5 and 6.7 m in a large multifloor building. CILoS also correctly differentiated between floors 90% of time. Letchner et al. used a hierarchical Bayesian network for wireless localization, and tested their model in a 7-floor building [73]. Krishnan et al. used fingerprinting with wireless sniffers for localization [46] which they tested in two floors. They used a heuristic majority-based logic to distinguish between floors. If a majority of the sniffers that "hear" the m strongest signals are in the same floor F and adding 5 dB to the $m+1$-st signal does not change this, then, the client is declared to be in floor F. The quantity m is called the decision depth. They reported that they could classify the floors correctly with $m \leq 3$, and in most cases $m = 1$ was sufficient.

To reduce the calibration effort (but not eliminate it completely), Wang et al. proposed a fingerprinting based localization algorithm [74] inspired by similar floor plans across floors in buildings. By collecting calibration data with known locations (labeled data) from a source floor, and with unknown locations (unlabeled data) from a target floor, they could localize mobile nodes on the target floor, i.e., labeled data can be used to train the localization system for all floors with only unlabeled data.

Techniques		Deployment		Calibration		Experiments & Results	
		Tech.	Maint. & Inst.	Locations/Area	Samples/Loc	Accuracy	Responsiveness
Calibration-Based	RADAR [3]	Wi-Fi	3 BSs in 43.5m x 22.5 m area	70 pts/54 rooms (F1) 20 pts/54 rooms 10 pts/54 rooms	4x20 4x20 4x20	2.94m,50th ~4m, 50th ~6m, 50th	>20sec per localization
	LEASE [18]	Wi-Fi	5 APs in 225x144ft^2	70 pts (F2) 10 pts		8 ft, median 15 ft, median	
	Neural Networks[1]	BT	3 BT dongles, 9m^2	19 pts (NN)		1.1 m median	
	JC [37]	Wi-Fi	12 APs in 224x85.1 ft^2	110 loc in corridors (JC)	300 samples/loc 1 sample/sec	3.2ft , median	
	IT [37]	Wi-Fi	12 APs in 224x85.1 ft^2	110 loc in corridors (IT)	300 samples/loc	3.8ft , median	
	Horus [28]	Wi-Fi	12 APs in 224x85.1 ft^2	110 loc in corridors (Horus)	300 samples/loc	2.15 ft mean	
	BI [38]	Wi-Fi	4APs 32.5m hallway	1.5m grids (BI)		0.3m, 50th 1.5m, 75th	
Calibration-Free							
	TIX [32]	Wi-Fi	4 APs 1020 sq. m.	207 pts (TIX)		5.4 m, mean	
	Redpin [34]	Wi-Fi, GSM, BT	15x26 sq. m.	26 rooms		90% true	
	Truncated SVD [31]	Wi-Fi	8 APs 23 m x 26 m	(SVD)		~3 m, mean	
	Wireless Sniffers [39]	Wi-Fi	3 sniffers, 1 AP in 160 sq. m.	160 pts (WS)		2.07m mean 1.84m mean	10 sec 100 sec
	Bayesian Graph Model [19]	Wi-Fi	5 APs in 225x144 ft^2	(NHBG): 20 pts (HBG): 100 pts		97.4ft 18.3ft 64ft 15.1ft	

Figure 2.16: Calibration-based and Calibration-free Localization Systems

3.0 SUB-AREA LOCALIZATION ALGORITHM

Coarse grained localization is usually less complex than fine grained localization that employs received signal strength (RSS) or time of arrival (ToA) or angle of arrival (AoA). A simple technique that falls into coarse grained localization category is proximity based localization which uses the connectivity information between mobile nodes (MNs) and monitoring stations (MoSs) to estimate a node's location.

Proximity based localization algorithms usually locate a MN at the same position as that of the closest point of association (e.g., access points - APs, RF-ID tags, or the MoS that receives signal with highest strength from the MN to be localized). For good accuracy and precision, such points of reference need to have a small range. In the extreme case, each reference point would cover a small area of radius r, and r would be the maximum localization error. According to the classical circle covering problem, [75], to cover an area with the smallest number of circles with radius r, the circles must be placed at the vertices of an equilateral triangle graph overlaying the target area. Each triangle has side $r\sqrt{3}$. However, this would require a large number of densely deployed reference points. In the next sections, we will show that the number of MoSs can be decreased by 41% compared to simply densely covering the required area with MoSs for the same maximum localization error, with a carefully chosen monitoring range [1]. Therefore, the coverage of reference points is as significant as the number and placement of reference points, and an effort to fine tune

[1] In a target area of size $L \times L$, the smallest number of vertices (circles) is

$$N = \left\lfloor \frac{\lceil 2(L-d)/3d \rceil + 1}{2} \right\rfloor \times \left(\left\lceil \frac{L}{d\sqrt{3}} \right\rceil + 1 \right) + \left\lceil \frac{\lceil 2(L-d)/3d \rceil + 1}{2} \right\rceil \times \left(\left\lceil \frac{L}{d\sqrt{3}} \right\rceil \right). \tag{3.1}$$

In the next sections, we will show that with 9 MoSs we can achieve a maximum localization error of $0.15L$ with SAL under ideal conditions. Covering the area with MoSs, each with range $0.15L$, requires at least $N = 22$ MoSs, which is more than twice this number.

these parameters may lead to achieve a given performance requirement in terms of a set of specified performance metrics.

Four performance metrics that are usually adopted in evaluation of localization performance are accuracy, precision, availability, and cost. The *accuracy* of a localization system denotes how close the estimated location is to the actual location and can be expressed in terms of the average localization error or the maximum localization error. *Precision* is the probability with which a given accuracy is achieved. *Availability* is the fraction of time and/or space that the location estimate is available to the user. And, the *cost* consists of the cost of installation and maintenance of the localization system, and it is directly proportional to the number of infrastructure entities used in a localization system. Typically, the accuracy is the metric of most importance, although this may be application dependent. Our initial focus will be on accuracy although we consider other metrics subsequently.

A previously proposed proximity based localization algorithm determines the estimated location as the centroid of the positions of MoSs that hear a MN. This algorithm is called ***Centroid Localization Algorithm (CA)*** [16]. We propose a novel idea of using the knowledge of the average monitoring ranges of MoSs and their intersections to improve the accuracy of localization. The intersection of monitoring ranges of MoSs create unique *sub-areas* that restrict the possibilities of where a MN is when multiple MoSs can hear the MN. We call this approach ***Sub-Area Localization (SAL)***.

In this chapter, we define a **localization scenario** as a localization algorithm using a given number (N) and placement ($\mathbf{P}$) of MoSs in a given localization area ($A = L \times L$), and denote this as $< N, \mathbf{P}, A >$. Then, we determine that for a given localization scenario, $< N, \mathbf{P}, A >$, the monitoring range of a MoS affects the localization performance significantly. Our objective in this chapter is to demonstrate the results of a broad investigation of the effect of monitoring range on localization performance under various localization scenarios and channel conditions.

1. In order to estimate the best monitoring range that minimizes the average localization error, we first approach this problem with an analytical framework using a scenario with grid placement of MoSs, $< N, \mathbf{G}, A >$, under ideal channel conditions, and show that the number, size and shape of sub-areas created by joint coverage areas can be used to

estimate the best monitoring range providing best accuracy. Furthermore, we show that with $< N, \mathbf{G}, A >$, the monitoring range providing the best accuracy falls into the range $(\sqrt{2}g/2, g)$, where g is the grid spacing. SAL can provide at least 30cm better accuracy compared to CA in a localization area of size $100m^2$. When the number of MoSs increases in this scenario, the grid spacing will get smaller, and intuitively this should reflect in a better accuracy. We show that the best monitoring range and corresponding accuracy follow an exponential decrease with exponentially increasing numbers of MoSs.

2. Localization performance may be different in different environments unlike the ideal case. In different environments, different channel conditions arise, and they have to be considered while determining the monitoring range that results in the best localization performance. We explore the effect of log-normal shadowing and wall attenuation factor (which incorporates RSS attenuation due to walls and other objects). Then, we show that irrespective of the placement scheme, under severe channel conditions, the best monitoring range approaches $\sqrt{2}g/2$ which is the range needed to barely cover the whole localization area.
3. In realistic localization environments, due to the presence of widely deployed wireless systems, there may be available APs that have been previously deployed, but can now be also used for localization. We explore the effect of incrementally adding MoSs to a SAL system, and show that increasing the number of MoSs can provide better accuracy. A formula generated by curve fitting demonstrates a 20cm improvement in a $10 \times 10\text{m}^2$ area with the addition of each MoS in a grid placement. Grid placement may not be possible due to various reasons such as walls, furnitures, etc. Therefore, we also show the effect of perturbed grid placement $< N, \tilde{\mathbf{G}}, A >$ on the best monitoring range and corresponding localization performance. As mentioned earlier the effort during installation and maintenance has a close relationship to the cost of the localization system. Therefore, we also investigate random placement schemes which require much less effort in deployment, and demonstrate that although random placement disturbs the number/shape and size of the subareas the accuracy is not affected significantly, and using SAL is still advantageous over the centroid localization algorithm.

This chapter is organized as follows. Sections 3.1 and 3.2 explain how SAL and CA works.

Then, in Section 3.3, we provide the analytical framework, and analyze the performance of SAL under ideal channel conditions. In Section 3.4, we compare SAL and CA performance under various channel conditions and explore the effect of multiple RSS thresholds and incremental addition of MoSs. Then, in Sections 3.5 and 3.6, we investigate the effect of perturbation of grid placement and random placement of MoSs.

3.1 SUB-AREA LOCALIZATION (SAL)

We first explain how SAL works in an ideal and general case. Let a number of MoSs be placed at known locations in a given area A with mean monitoring range d. If a transmission occurs from a MN that is at a distance less than or equal to d from a MoS, the MoS can "hear" the MN – otherwise it cannot. Ideally, the monitoring range is a circular area with radius d for each MoS. This model resembles the unit disk graph model commonly used in the ad hoc networking literature. According to this standard path-loss model [76], the power of the received signal at the MoS, P_r, can be calculated as,

$$P_r = P_t - 10\alpha \log(d_{ij}) + X, \tag{3.2}$$

where P_t is the MN's transmit power, α is the path-loss coefficient, d_{ij} is the distance between MoS_i and MN_j, and $X = N(0, \sigma)$ is a normal random variable with zero mean and standard deviation σ representing the effect of shadowing on the received power. Under ideal channel conditions $X = N(0, 0)$. A MN at distance d_{ij} can be detected if P_r is equal to at least the detection threshold of the MoS, R_{th} (R_{th} must be at least at the receiver sensitivity level).

$$\begin{aligned} P_r &\geq R_{th} \\ P_t - 10\alpha \log(d_{ij}) &\geq P_t - 10\alpha \log(d) \\ -10\alpha \log(d_{ij}) &\geq -10\alpha \log(d). \end{aligned} \tag{3.3}$$

We assume that P_t is known by the localization system and fixed for all mobile devices; and R_{th} can be tuned so that the required monitoring range, d, can be achieved.

Then, under ideal conditions, we can partition the area A into K distinct *small* subareas $A_i, i = 1, 2, \cdots, K$, i.e., $A = \bigcup_{i=1}^{K} A_i$, and $|A| = \sum_{i=1}^{K} |A_i|$ such that each of these subareas is monitored by a unique subset of MoSs, and area A can be wholly covered by MoSs. For example consider Figure 3.1. Here, four MoSs of range $d = L$ are placed at the corners of an $L \times L$ network and they create $K = 9$ unique subareas. When MoSs detect a MN, they could report this information along with the time of detection to a central server. With the knowledge of which MoSs heard the MN and *which did not*, the correct subarea can be identified and the estimated location is the center of mass of the identified subarea. In Figure 3.1, when all MoSs detect a transmission from the MN, the subarea is A_5, and the estimated location is the center of this area, shown as point P in the figure.

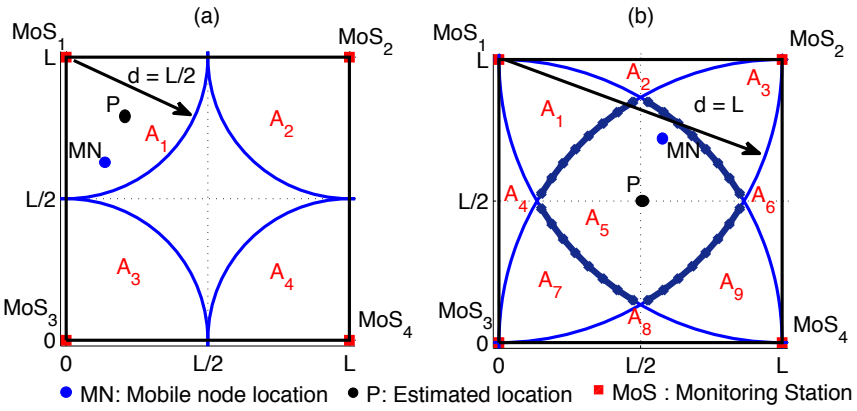

Figure 3.1: The sub-area localization

We assume that there can exist only one transmission at a time in a subarea covered by a set of MoSs, and a MN detected by MoSs must be in area A.

3.1.1 Accuracy

We employ error metrics defined as follows in our analyses and evaluations in later sections. Let us suppose that when N MoSs are deployed, there exist K distinct subareas, A_i, $i = 1, 2, \ldots, K$, where different subsets of MoSs can hear a MN. Let point $P_i = (\tilde{x}_i, \tilde{y}_i)$ represent

the coordinates of the center of mass (centroid) of area A_i, which is also the estimated location of any MN detected in area A_i. As there is no closed-form solution for the centroid of a subarea, we assume a large number M of MNs are placed uniformly randomly in A with M_i nodes in subarea A_i. Then, we calculate the coordinates of point P_i as,

$$(\tilde{x}_i,\ \tilde{y}_i) = \left(\frac{\sum_{\forall x \in A_i} x}{M_i},\ \frac{\sum_{\forall y \in A_i} y}{M_i} \right). \tag{3.4}$$

The *maximum localization error* of a MN detected in subarea A_i is equal to the distance between P_i and the point which is the farthest away from P_i in A_i. We denote the maximum localization error in subarea A_i by $\epsilon_{max}(A_i)$:

$$\epsilon_{max}(A_i) = max_{x,y \in A_i} dist[(x,y),(\tilde{x}_i,\tilde{y}_i)], \tag{3.5}$$

where (x,y) is any point in subarea A_i, and $dist[(x,y),\ (\tilde{x}_i,\tilde{y}_i)]$ is the Euclidean distance between (x,y) and $P_i = (\tilde{x}_i,\tilde{y}_i)$. Then, the average maximum localization error is,

$$E_{max} = \sum_{i=1}^{K} P(MN\ in\ A_i)\epsilon_{max}(A_i), \tag{3.6}$$

where $P(MN\ in\ A_i)$ is the probability that a MN is in subarea A_i. If all locations in A are equally likely for any MN,

$$P(MN\ in\ A_i) = |A_i|/|A| = M_i/M. \tag{3.7}$$

Similarly, the *mean localization error* of a MN detected in subarea A_i is equal to the average distance between P_i and all of the points (locations) in subarea A_i. The mean localization error in area A_i and the overall mean localization error, respectively, are,

$$\epsilon_{avg}(A_i) = \frac{1}{M_i} \sum_{\forall (x,y) \in A_i} dist[(x,y),\ (\tilde{x}_i,\tilde{y}_i)], \tag{3.8}$$

$$\begin{aligned} E_{avg} &= \sum_{i=1}^{K} P(MN\ in\ A_i)\epsilon_{avg}(A_i) \\ &= \frac{1}{\sum_{i=1}^{K} M_i} \sum_{i=1}^{K} \sum_{\forall (x,y) \in A_i} dist[(x,y),(\tilde{x}_i,\tilde{y}_i)]. \end{aligned} \tag{3.9}$$

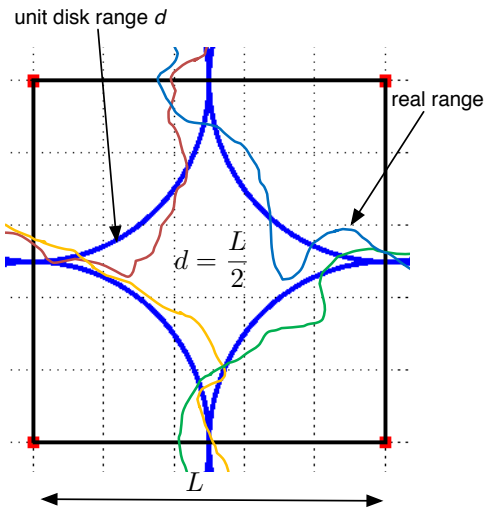

Figure 3.2: Real and unit disk ranges

We compute the maximum and mean localization errors numerically as there are no easy closed form solutions. It is also possible to compute the *cumulative distributions* of errors assuming that a MN's location can be anywhere in the area. Note that maximum localization error is the accuracy that can be achieved in a subarea with 100% precision, whereas the median localization error is the accuracy with 50% precision.

3.1.2 Availability

The monitoring range, d, can vary with time since it depends on the path-loss and radio propagation in addition to the transmission power of MNs and the sensitivity of MoSs, i.e., $X = N(0, \sigma),\ \sigma > 0$ in Eqn. 3.2. Therefore, the actual coverage area of a MoS may not be an exact circle as shown in Figure 3.2. We observe that the joint areas between MoSs may not be symmetric, and in fact, under certain circumstances joint areas between groups of MoSs may emerge while they do not exist assuming fixed circular ranges. In SAL, MoSs do not collect samples from known locations to capture channel parameters either in an offline or in an online phase as in [2] and [58]. A MN can be detected by a set of MoSs that have a subarea as defined in Section 3.1 assuming a unit disk model. Therefore, it can be detected

by a set of MoSs that do not create a subarea if the unit disk model is used, but in reality do have intersecting monitoring areas because of radio propagation peculiarities. For example, with 9 MoSs in grid placement, when $d < \sqrt{2}L/2$, the monitoring ranges of MoS_1 at (0,0) and MoS_9 at (L, L) are not supposed to intersect if unit disk model is used. However, the monitoring system may still observe that a MN is heard by both MoS_1 and MoS_9 (i.e., they have an intersection). A MN may not be detected by any MoSs at all because there are coverage holes not identified by the unit disk model. There is no way that the localization system can even know about the existence of these MNs; therefore, their locations are not estimated (i.e., the coverage is not 100%).

We define localization availability as the fraction of area in A that the location estimate is available to the user. We are interested in availability in space rather than time, and we assume that the availability is averaged out in time. Let C' be the size of the area in A that can be sensed by at least one MoS. When $\sigma = 0$, MNs are always located in valid sub-areas whose boundaries are determined from the unit disk model. When $\sigma > 0$, an area $C_{invalid}$ of the localizable area C' may belong to invalid subareas. These areas arise because radio propagation vagaries may create intersections of monitoring areas of MoSs, that theoretically do not have intersecting ranges. SAL cannot be directly used for localizing nodes in these invalid subareas. Nodes in such invalid subareas can be localized for instance, by using the centroid localization algorithm. Let $C = C' - C_{invalid}$. Then, the fraction of the area in which a node can be localized by SAL is given as the availability of SAL (R_{SAL}),

$$R_{SAL} = \frac{C}{|A|} \tag{3.10}$$

Under ideal conditions, and given that the whole localization area is covered by the monitoring ranges of MoSs, i.e., $d \geq \sqrt{2}g/2$, $C = C' = |A|$; the availability of SAL is 100%. However, channel conditions may cause deformations on monitoring ranges, then, availability may be affected negatively. We show the effect of these differences on the localization error using simulations with a normally distributed shadowing factor (in dB) in Section 3.4.

3.1.3 Efficiency

The efficiency definition for a localization system depends on the requirements of localization system application. The application may require location estimates that are as accurate as possible when they are available; or it may require estimates in as large an area as possible without caring a lot about the accuracy. Another parameter may be the cost; we may need to provide the location estimate using a small budget by trading off accuracy or availability; or we may need to determine the cost for a given requirement of target $< E_{avg},\ R >$ pair. When we bring together accuracy, availability and the number of monitoring stations as a representative of system cost, we define an efficiency metric, which we call localization efficiency, η. This metric is directly proportional to R and reversely proportional to E_{avg} and N. Thus, the localization system is more efficient when the localization error and the number of infrastructure entities used are small, and availability factor is large. Efficiency metric is given as,

$$\eta_{SAL}(i,j) = \frac{R_{SAL}}{N^i \times (E_{avg,SAL}/L)^j}, \tag{3.11}$$

where the impact of the localization error and the cost in terms of the number of MoSs deployed weighed with i and j, respectively.

3.2 CENTROID ALGORITHM (CA)

Instead of using the sub-areas, it is possible to use the centroid of the locations of the MoSs that hear the nodes as the estimated location of a MN. This algorithm is proposed in [16]. Again, the monitoring range d influences which MoSs hear the nodes and thus what the error is, in the location estimate. The difference here is that the estimated location is *not* the center of mass of the subareas (the subareas can be small) but the centroid of the MoS locations. Let $\overline{MoS}$ be the set of MoSs that detects a MN, and (x_i, y_i) be the coordinates of MoS_i, then the estimated location P_i is given as,

$$(\tilde{x}_i,\ \tilde{y}_i) = \left(\frac{\sum_{\forall MoS_i \in \overline{MoS}} x_i}{|\overline{MoS}|},\ \frac{\sum_{\forall MoS_i \in \overline{MoS}} y_i}{|\overline{MoS}|} \right), \tag{3.12}$$

Table 3.1: SAL acronyms and variables

Variable	Explanation
$A = (L \times L)$	Localization area
N	Number of Monitoring Stations
$< N, \mathbf{P}, A >$	Localization algorithm using N MoSs, with placement scheme $\mathbf{P}$ in area A
d	Monitoring range
$A_i,\ i = 1, 2, \ldots, K$	Unique subarea
R_i and k_i	Region and Number of occurance of the same region
D_i	Diagonal of subarea A_i
$P_i(x_i, y_i)$	Two dimensional coordinates of the center of mass of subarea A_i
$X = N(0, \sigma)$	Log normal shadowing with variance σ^2
E_{avg}	Expected average localization error
E_{max}	Expected maximum localization error
R	Availability of localization algorithm
η	Efficiency of localization algorithm

The availability of location estimates with CA only depends on whether the MN is detected by at least one MoS. In that case, the estimated location is the location of that MoS and availability is 100%. Then, the fraction of MNs that can get location estimate from

CA is given as the availability of CA (R_{CA}),

$$R_{CA} = \frac{C}{|A|} = \frac{C'}{|A|} \tag{3.13}$$

In a manner similar to the definition in SAL, localization efficiency of CA is given as,

$$\eta_{CA}(i,j) = \frac{R_{CA}}{N^i \times (E_{avg,CA}/L)^j}. \tag{3.14}$$

3.3 PERFORMANCE UNDER IDEAL CHANNEL CONDITIONS

From Eqs. 3.6 and 3.9 in Section 3.1, it is clear that the accuracy of SAL depends on the size ($|A_i|$) and number (K) of subareas which depend on the number (N) and placement ($\mathbf{P}$) of MoSs as well as the size and shape of localization service area (A). For a given localization scenario $< N, \mathbf{P}, A >$, the partitioning of the area into subareas, and therefore, the ***number*** of subareas (K) and their ***sizes*** ($|A_i|$) change with the monitoring range d. Note that in Figure 3.3 when $d = L/2$, the center of the network is not covered by the MoSs. In this case, a MN in this area cannot be localized. As the monitoring range d increases, the network gets completely covered. As d continues to increase, at some point, each MoS can hear a MN where ever it is in the network and the estimated location will always be at the center of the network. There is an optimal d where the network is covered and the (average or maximum) localization error is the smallest. Therefore, the monitoring range d is a parameter that has a significant effect on the accuracy and precision (and eventually availability).

In this section, we investigate the effect of monitoring range (d) on accuracy that can be achieved by SAL under ideal channel conditions (unit disk model). However, we are aware that this is not a realistic assumption and the real monitoring contour depends on the transmission powers of nodes, the environment, and the localization protocol. We use this assumption to develop a tractable analysis and to obtain insights on the problem. We note that under ideal channel conditions, availability is 100% when $d \geq \sqrt{2}g/2$; therefore, only accuracy and precision affect the efficiency of localization algorithm.

3.3.1 Subarea Characterization

For our analysis, we first assume that a given number of monitoring stations (N) are placed on a virtual grid with grid spacing $g = \frac{L}{\sqrt{N}-1}$ in a square shaped area A of size $L \times L$ as in localization scenario $< N, \mathbf{G}, A >$. Several instances of joint subareas of $< 4, \mathbf{G}, A >$ and $< 9, \mathbf{G}, A >$ scenarios for different d values are shown in Figures 3.3 and 3.4, respectively (the overall area A is the black rectangle – a larger area is shown to illustrate the monitoring ranges/areas). From Figures 3.3 and 3.4, we observe that the number, shape and sizes of subareas change as monitoring range changes. Tables 3.2 and 3.3 show the *number* of subareas, K, for various ranges of d with 4 and 9 MoSs.

We note that for each case, the *sizes* and *shapes* of the subareas vary as d changes in the specified range. Therefore, there is space to optimize the accuracy by tuning the monitoring range. Intuitively, we expect the accuracy to be better with larger K and smaller $|A_i|$. An upper bound on K (irrespective of d) is that $K \leq 2^N - 1$, where there are N MoSs. This is because there could potentially be subareas that are covered by exactly one MoS, by combinations of two MoSs, three MoSs and so on for a total of $\sum_{i=1}^{N} \binom{N}{i} = 2^N - 1$ subareas. However, under ideal channel conditions, not all such combinations are possible because we assume that each MoS has the same range. Another way of looking at the subareas is to think of them as RSS fingerprints, but quantized to such an extent that only the presence or absence of an RSS value is observed (i.e., a MoS can hear the MN or not).

Considering only the number of subareas, Case D with $N = 4$ (shown in Figure 3.5) and Case G with $N = 9$ represent the best intervals for the best monitoring range. With $N = 4$, the upper bound on K is 15, but it is physically impossible to have only the diagonally opposite MoSs cover a subarea (i.e., subareas covered by exactly MoS_1 and MoS_4 or MoS_2 and MoS_3 in Figure 3.5) when all MoSs have the same range. We note that the lower bound on the best monitoring range is also the range that is providing 100% availability under ideal channel conditions.

With a given monitoring range, there may exist many small areas and a smaller number of larger areas which will cause higher overall error. As previously discussed, we would like to have several small regions for good localization accuracy. Due to the symmetry in placement

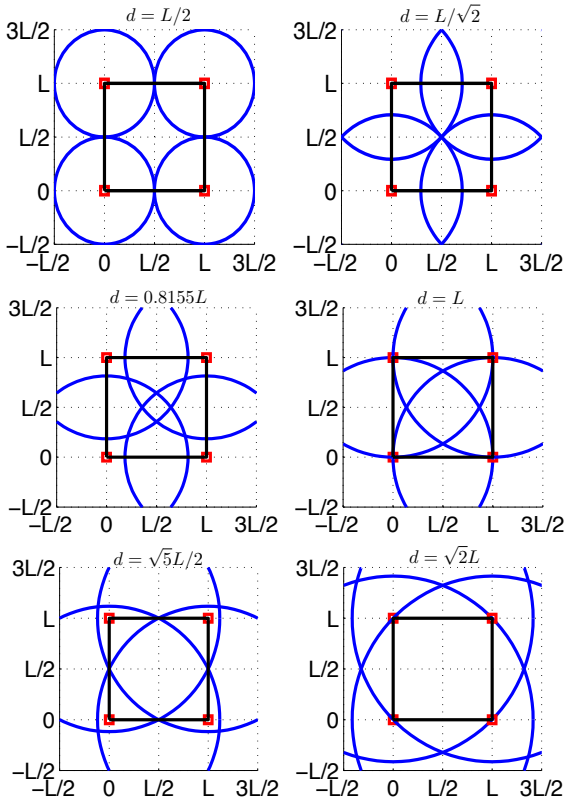

Figure 3.3: Visualization of subareas in localization scenario $< 4, \mathbf{G}, A >$

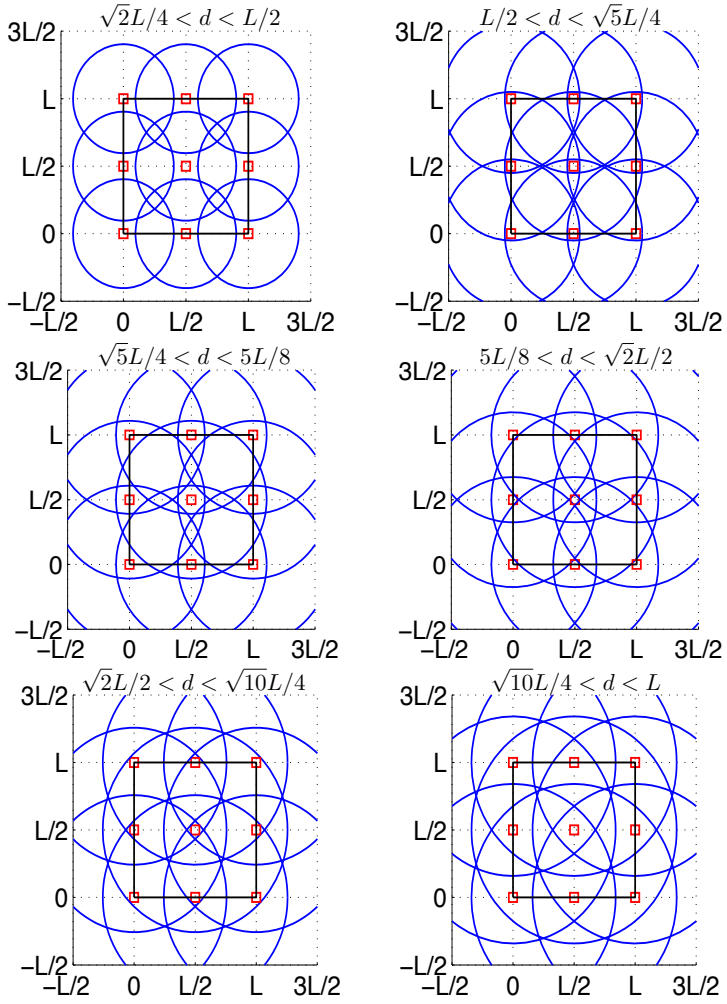

Figure 3.4: Visualization of subareas in localization scenario $< 9, \mathbf{G}, A >$

and ranges of MoSs, subareas can be classified into a smaller set of types of "regions". For example, four distinct types of regions with $N = 4$ MoSs are shown in Figure 3.5. To take into account both the number of subareas and the size of distinct regions into consideration, we evaluate the ratio of the number of subareas K to the maximum "region" size :

$$max\left(\frac{K}{max(R_i)}\right) \tag{3.15}$$

In Figure 3.6, we plot the ratio of K and the maximum area among all R_i's for a given d with 9 MoSs using numerical computation. From this figure, clearly, the best monitoring

Table 3.2: Number of distinct subareas A_i as a function of d; $N = 4$ MoSs

Case	Interval	**K**
A	$0 < d \leq L/2$	4
B	$L/2 < d \leq \sqrt{2}L/2$	9
C	$d = \sqrt{2}L/2$	8
D	$\sqrt{2}L/2 < d < L$	13
E	$d = L$	9

range with $N = 9$ falls into the interval defined by Case C. An instance of case D and case C for $N = 4$ and $N = 9$ are shown in Figures 3.5 and 3.7. Note that, Case C with $N = 9$ is a cascaded version of Case D with $N = 4$. For larger N, the visualization of the subareas in the best case will remain the same. Then, for a localization scenario with grid placement $< N, \mathbf{G}, A >$ with grid spacing $g = \frac{L}{\sqrt{N}-1}$, the best monitoring range minimizing the localization error is in the interval $\sqrt{2}g/2 < d < g$.

3.3.2 Heuristics Towards the Best Monitoring Range

In the previous section, we showed that with grid placement, and under ideal channel conditions, the range $\sqrt{2}g/2 < d < g$ represents a potential interval for monitoring range which can provide the best accuracy. In this section, we introduce four heuristics towards best monitoring range that we evaluate in this interval. Two of these heuristics depend on the areas of the distinct regions $R_i(d)$ as they change with d and the number of such regions $k_{i,d}$, whereas the remaining two consider the shapes of these distinct regions instead of sizes.

To explain these heuristics and to evaluate them later, we start with computing the *sizes* of the subareas with $N = 4$. In case D, the 13 distinct subareas can be classified into 4 types of regions, R_i, $i = 1, 2, 3, 4$ as shown in Figure 3.5. If C_j is the monitoring area of

Table 3.3: Number of distinct subareas A_i as a function of d; $N = 9$ MoSs

Case	Interval	**K**
A	$d = L/4$	9
B	$L/4 < d \leq \sqrt{2}L/4$	21
C	$\sqrt{2}L/4 < d < L/2$	41
D	$d = L/2$	32
E	$L/2 \leq d < \sqrt{5}L/4$	41
F	$d = \sqrt{5}L/4$	29
G	$\sqrt{5}L/4 < d < 5L/8$	49
H	$d = 5L/8$	37
I	$5L/8 < d < \sqrt{2}L/2$	41
J	$d = \sqrt{2}L/2$	32
K	$\sqrt{2}L/2 < d < \sqrt{10}L/4$	37
L	$d = \sqrt{10}L/4$	25
M	$\sqrt{10}L/4 < d \leq L$	37

MoS_j, the centers of C_1, C_2 and C_4 are located at $(0, L)$, (L, L) and $(L, 0)$, respectively. The calculation of size of regions R_i based on the intersection areas $B_1(d, g) = C_1 \cap C_2$ and $B_2(d, g) = C_1 \cap C_4$ as functions of d and g are given as:

$$B_1(d, g) = 2(R_2 + 2R_3 + R_4) = 4\left[\frac{\pi d^2}{2\pi/\arccos(g/2d)} - \frac{\sqrt{d^2 - g^2/4}}{4}\right] \tag{3.16}$$

$$B_2(d, g) = 2R_3 + R_4 = 4\left[\frac{\pi d^2}{2\pi/\arccos(\sqrt{2}g/2d)} - \frac{\sqrt{d^2 - g^2/2}}{4}\right] \tag{3.17}$$

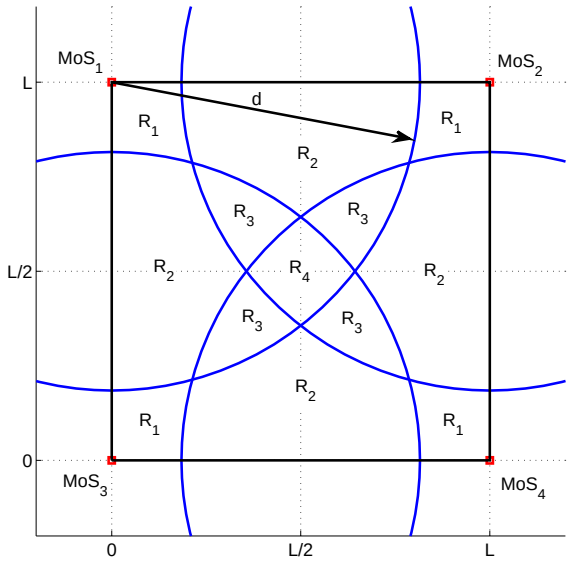

Figure 3.5: Visualization of regions in A when $\sqrt{2}L/2 < d < L$

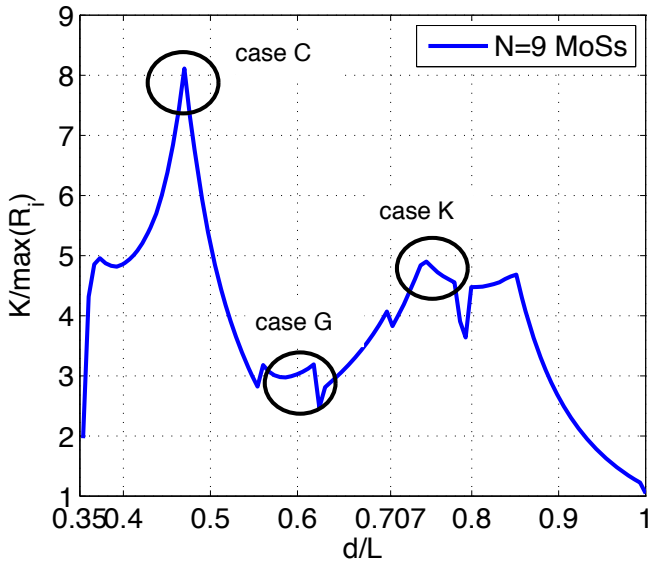

Figure 3.6: Ratio of K to the maximum area among all R_i's for various d/L

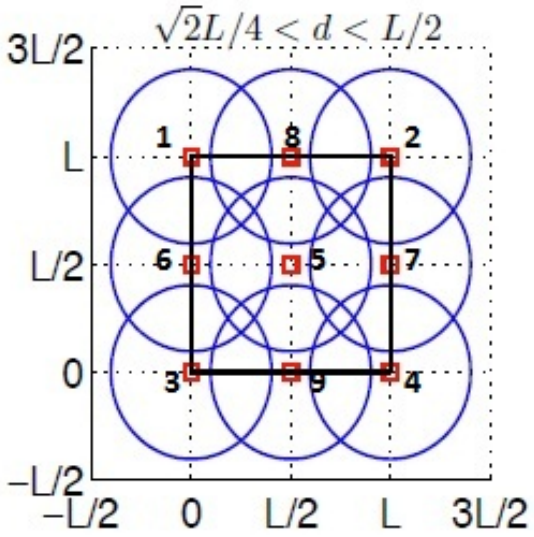

Figure 3.7: Visualization of subareas when $N = 9$, $\sqrt{2}L/4 < d < L/2$

Other relations between R_i and L can be derived as follows:

$$4(R_1 + R_2 + R_3) + R_4 = L^2 \tag{3.18}$$

$$R_1 + 2R_2 + 3R_3 + R_4 = \frac{\pi d^2}{4}. \tag{3.19}$$

Then the areas of regions $R_i,\ i = 1, 2, 3, 4$ can be calculated as,

$$R_1 = \frac{B_2(d, g) + g^2}{2} + \frac{\pi d^2}{4} \tag{3.20}$$

$$R_2 = \frac{B_1(d, g)}{2} - B_2(d, g) \tag{3.21}$$

$$R_3 = \frac{B_2 - g^2 + \pi d^2 - 2B_1(d, g)}{2} \tag{3.22}$$

$$R_4 = g^2 - \pi d^2 + 2B_1(d, g) \tag{3.23}$$

Figure 3.8 shows how the areas of the four types of regions R_i change with increasing d when $\sqrt{2}L/2 < d < L$. The y-axis in this plot corresponds to the percentage of area in A that a region R_i occupies, i.e., $R_i \times 100/|A|$. From Figure 3.8, we see that regions R_1, R_3 and R_4 are small when d is roughly $0.8L$. As previously discussed, we would like to have several small regions for good localization accuracy. We will see later that this visual observation

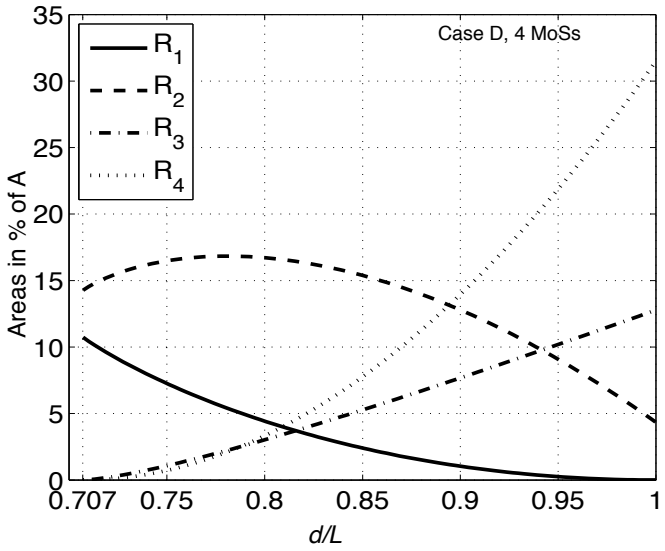

Figure 3.8: Areas of regions R_i in % of A as a function of d/L, $N = 4$

of the potentially best value of d for good accuracy is confirmed through the numerically computed error metrics.

When $N = 9$, there are 7 types of regions R_i in case C that make up the 41 distinct subareas. Once again, we note that case C with $N = 9$ is basically the concatenation or cascade of case D with $N = 4$ MoSs. Thus, we can compute the areas of these regions in a manner similar to the method used for $N = 4$ case. Four of these subareas can be calculated from equations 3.20-3.23, and the remaining three are:

$$R_5 = 2R_1; \quad R_6 = 2R_2; \quad R_7 = 4R_1. \tag{3.24}$$

Figure 3.9 shows how the areas of the seven types of regions R_i change with increasing d when $\sqrt{2}g/2 < d < g$. From Figure 3.9, we see that regions are small when d is roughly $0.425L$. As previously discussed, we would like to have several small regions for good localization accuracy. Again, we will see later that this visual observation of the potentially best value of d for good accuracy is confirmed through the numerically computed error metrics.

Now we define the four heuristics. The quantity $\hat{d}_1$ gives the d that minimizes the ratio between the largest region and the smallest region created over $\sqrt{2}g/2 < d < g$; the quantity

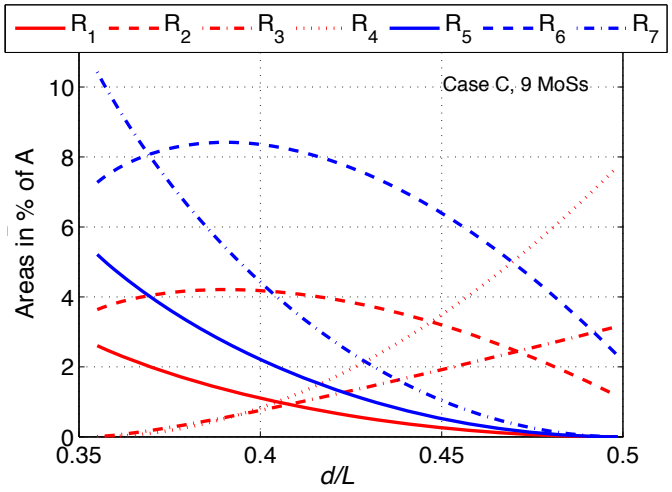

Figure 3.9: Areas of regions R_i in % of A as a function of d/L, $N = 9$

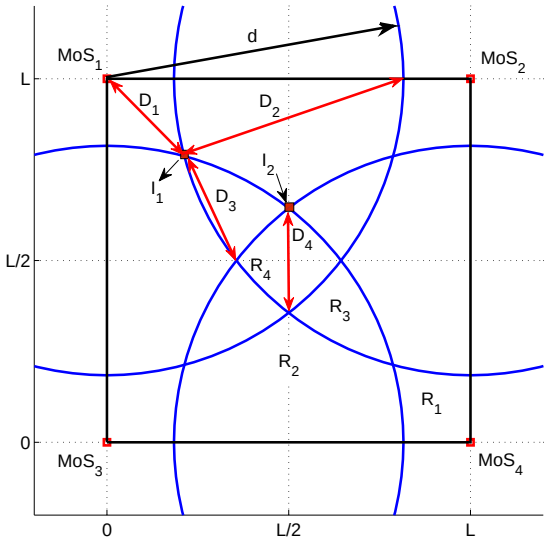

Figure 3.10: Visualization of diagonals and intersection points in A when $\sqrt{2}L/2 < d < L$

$\hat{d}_2$ weighs this ratio by the *number* of regions – for example, there are $k_{2,d} = 4$ regions R_2 in case D in Table 3.2. The two other heuristics consider the shapes of regions. The intuition for these two heuristics comes from comparing the maximum error in a square and

a rectangle of the same size. When we think of a square of size $a \times a = |A|$, and a rectangle of size $\frac{a}{2} \times 2a = |A|$, although the area of these two shapes are the same, the maximum error in the square is $\sqrt{2}a/2$, whereas it is $\sqrt{17}a/2$ in the rectangle. The maximum error here is calculated by the largest possible distance between two MNs in a region divided by two, which corresponds to the diagonal size divided by two. In other words, although two regions may have the same area, their shapes may contribute to very different localization errors. Therefore, the two other heuristics ($\hat{d}_3$ and $\hat{d}_4$) are based on the length of the diagonals of regions, $D_i(d)$ as shown in Figure 3.10. We define the diagonal as the largest possible distance between two MNs in subarea A_i. The four different diagonal lengths when $N = 4$ are given as,

$$D_1(d) = dist[(0, g), I_1] \tag{3.25}$$

$$D_2(d) = dist[(d, g), I_1] \tag{3.26}$$

$$D_3(d) = dist[I_1, I_2] \tag{3.27}$$

$$D_4(d) = 2I_2(y) - g, \tag{3.28}$$

where $g = L$ when $N = 4$, and I_1 and I_2 are the intersection coordinates of MoS_2 and MoS_3, and MoS_3 and MoS_4, respectively. These intersection coordinates are calculated as follows and shown on Figure 3.10.

$$I_1(d) = \left(g - \sqrt{\frac{d^2}{1 + m_1^2}},\ g - \frac{m_1 d}{\sqrt{1 + m_1^2}} \right), \tag{3.29}$$

$$m_1 = tan\left(\frac{\pi}{4} - acos\left(\frac{g}{\sqrt{2}d} \right) \right),$$

$$I_2(d) = \left(g - \sqrt{\frac{d^2}{1 + m_2^2}},\ \frac{-m_2 d}{\sqrt{1 + m_2^2}} \right), \tag{3.30}$$

$$m_2 = -tan\left(acos\left(\frac{g}{2d} \right) \right)$$

For $N = 9$, 3 more distinct regions are created and the diagonals for these regions can be calculated as follows:

$$D_5(d) = dist[(0, 2g - d), I_1];\ \ D_6(d) = 2(g - I_2(y));\ \ D_7(d) = 2D_1(d). \tag{3.31}$$

The third heuristic quantity $\hat{d}_3$ is the monitoring range that minimizes the weighted sum of the diagonals and the last heuristic quantity $\hat{d}_4$ is the monitoring range that minimizes the difference between the largest and smallest diagonals (i.e., it tries to balance the shapes of the regions). Table 3.4 summarizes our findings from these four heuristics. The smallest estimate for best monitoring range is given by the 3rd heuristic. We observe differences in the effect of different heuristics when $N = 4$ and $N = 9$. The reason is that with 9 MoSs three more distinct regions are created. We will refer to this table in the next sections when comparing best monitoring ranges calculated with simulations to ranges calculated from these heuristics.

Table 3.4: Heuristics for $\hat{d}_i,\ i = 1, 2, 3$ for $N = 4$ and $N = 9$

Heuristic	4 MoSs	9 MoSs
$\hat{d}_1 = argmin_d \left[\frac{\max_i(R_i(d))}{\min_i(R_i(d))}\right]$	$0.815L$	$0.407L$
$\hat{d}_2 = argmin_d \left[\frac{\max_i(R_i(d)k_{i,d})}{\min_i(R_i(d)k_{i,d})}\right]$	$0.859L$	$0.405L$
$\hat{d}_3 = argmin_d \left[\sum_i D_i(d)k_{i,d}\right]$	$0.748L$	$0.356L$
$\hat{d}_4 = argmin_d \left[\max_i(D_i(d)) - \min_i(D_i(d))\right]$	$0.783L$	$0.405L$

It is also possible that from D_i, we can estimate the accuracy with 100% precision approximately as follows:

$$\tilde{E}_{max}(d) = \frac{1}{K}\sum_{i=1}^{K} k_{i,d}\frac{D_i(d)}{2}. \tag{3.32}$$

Note that this equation resembles the heuristic $\hat{d}_3$; therefore, the minimum $\tilde{E}_{max}$ is $0.202L$ with $N = 4$ at $d = 0.748L$, and $0.08L$ with $N = 9$ at $d = 0.356L$.

3.3.3 Evaluation of Localization Error

In order to obtain a more concrete picture, we numerically evaluated the error metrics defined in Section 3.1.1 using MATLAB. We assume N MoSs with same monitoring range d are placed on a grid in area A of size $L \times L$ as in scenario $< N, \mathbf{G}, A >$. Given scenario $< N, \mathbf{G}, A >$, $10^4 L^2$ MNs were randomly placed in A using a uniform distribution. The locations of MNs that are inside a subarea A_i are determined, and the center of mass of those MNs, calculated as $(\tilde{x}_i,\ \tilde{y}_i)$, is defined as the estimated location for each of those MNs in subarea A_i as explained in Section 3.1.

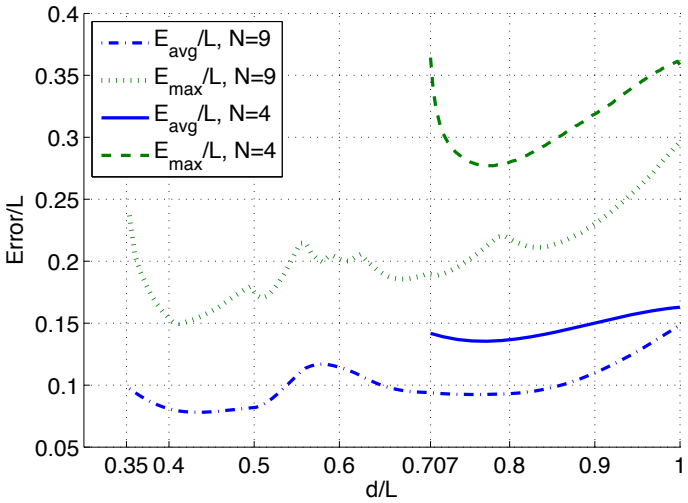

Figure 3.11: Average and maximum error in $< 4, \mathbf{G}, A >$ and $< 9, \mathbf{G}, A >$

We ran exhaustive simulations to evaluate $E_{max},\ E_{avg},\ \epsilon_{max},\ \epsilon_{avg}$, for different values of L and d with $N = 4$ MoSs and $N = 9$ MoSs. Figure 3.11 shows E_{max}/L and E_{avg}/L as a function of d/L for $d \geq \sqrt{2}g/2$ for 100% coverage of area A. With $N = 4$ MoSs, a monitoring range of $d = 0.77L$ provides the smallest mean localization error of 14% of one side of the unit of area $A = L \times L$. Similarly, the smallest mean accuracy of 8% of one side of the area A can be achieved when $N = 9$ with $d = 0.41L$. We observe that optimal d providing the minimum error falls into the range $\frac{\sqrt{2}\ g}{2} < d < g$ determined in Section 3.3 for both $N = 4$ MoSs and $N = 9$. We also observe increases and decreases in E_{max} and E_{avg} when $N = 9$ as d increases. This is due to the changing numbers and sizes of subareas for different values

of d. Also, we can observe that higher $\frac{K}{max(R_i)}$ regions in Figure 3.6 match the lower error regions in Figure 3.11.

Table 3.5 provides the maximum and average error and corresponding best monitoring ranges with 4 and 9 MoSs. When we compare the optimal d values in Table 3.4 and Table 3.5, the first and fourth heuristics, $\hat{d}_1$ and $\hat{d}_4$, provide the closest results to optimal d according to numerical calculations. And $\tilde{E}_{max}$ in Eq. 3.32 provides an upper limit on the accuracy.

Figure 3.12 shows the CDF of the localization error for the d values that minimize the mean localization error for $N = 4$ and $N = 9$ MoSs that can be used for additional analysis. For instance, from the CDFs, we see that SAL can provide an accuracy of $0.17L$ and $0.1L$ with a precision of 70% with $N = 4$ and $N = 9$ MoSs respectively.

Table 3.5: Optimal monitoring range $\hat{d}$ and minimum localization error

	4 MoSs	9 MoSs
$min_d[E_{max}]$	$0.28L$	$0.15L$
$\hat{d} = argmin_d[E_{max}]$	$0.765L$	$0.411L$
$min_d[E_{avg}]$	$0.14L$	$0.08L$
$\hat{d} = argmin_d[E_{avg}]$	$0.768L$	$0.431L$

We showed that the optimal monitoring range when $N = 9$ falls into the case C which is realized by concatenation of subareas created by $N = 4$ in case D. Therefore, the pattern of subareas repeats for larger N. It is possible to confirm that the above analysis can be generalized for $N = \left(\frac{L}{g} + 1\right)^2$ MoSs when N MoSs are on a grid with grid spacing $g = \frac{L}{\sqrt{N}-1}$. As Figure 3.13 for $N = 25$ MoSs shows, the pattern of subareas repeats for larger N. The best values for $\hat{d}$ fall in the range $\frac{\sqrt{2}}{2} g < d < g$. We ran simulations (not shown here) to confirm this. A curve fitting model indicates that the best values for d and the corresponding minimum localization error follow an exponential model. Figures 3.14 and 3.15 show the smallest E_{max} and E_{avg} values for different numbers N of MoSs and the curve fits that are obtained with MATLAB. The smallest E_{max} with $N = 4$ MoSs is $0.15L$, whereas it is less

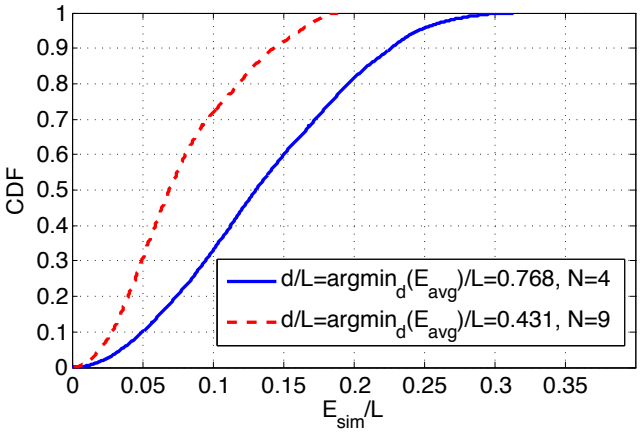

Figure 3.12: CDF of localization error

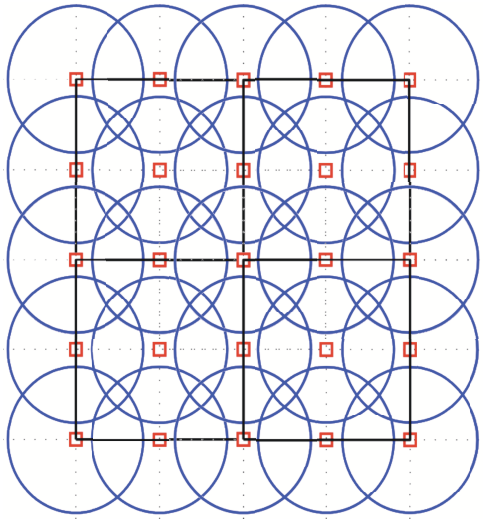

Figure 3.13: Subareas for smallest localization error with $N = 25$ MoSs

than $0.05L$ with $N = 25$ MoSs. Depending on the localization error requirement, having a large number of MoSs is appropriate for larger network sizes (e.g., an accuracy of 5m can be achieved in an area of size 10000m^2 with 25 MoSs).

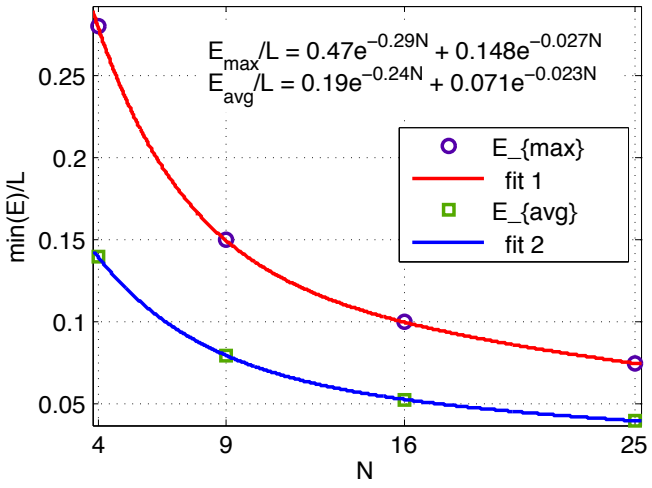

Figure 3.14: The smallest mean and maximum localization errors vs. N and curve fits

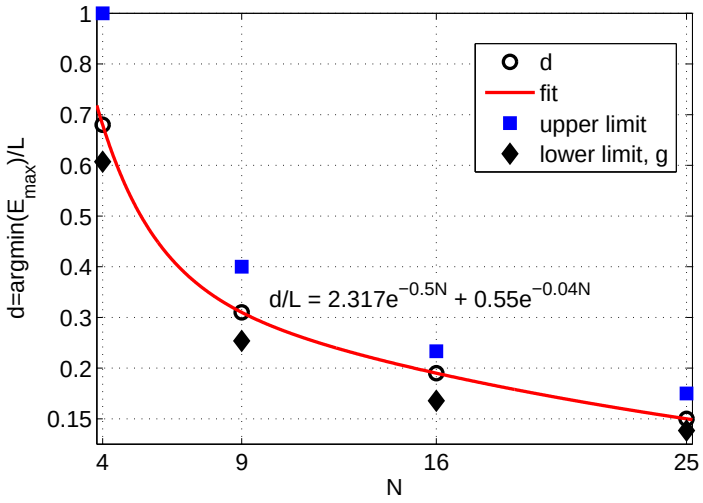

Figure 3.15: Best values of d (that minimizes the maximum error) vs. N

3.3.4 Evaluation of Efficiency

Having a large number of MoSs can provide better accuracy even for larger network sizes, but could be costly. In Figure 3.16, we plot the efficiency metric, in which the impact of the

localization error and the cost in terms of the number of MoSs deployed weighed with i and j, respectively:

$$\eta(i,j) = \frac{1}{N^i (E/L)^j} \tag{3.33}$$

Note that since $d \geq \sqrt{2}g/2$, availability is 100%, $R = 1$.

In this figure, η_{max} is calculated with E_{max}, whereas η_{avg} is calculated with E_{avg}. When accuracy and cost are weighed equally ($i = j = 1$), η decreases until the number of MoSs is 36, and then starts to increase with increasing number of MoSs. It appears that using smaller numbers of MoSs is better if both cost and localization error are equally weighed. The rest of the results are in line with intuition: if $i > j$, i.e., $i = 2,\ j = 1$, then, using smaller number of MoSs is more efficient, whereas if $j > i$, i.e., $i = 1,\ j = 2$, using higher number of MoSs is more efficient.

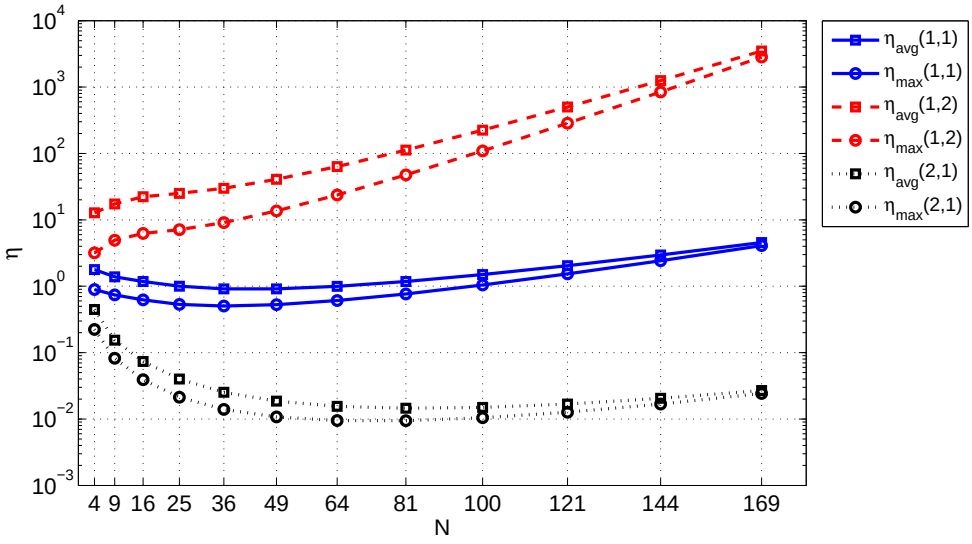

Figure 3.16: Efficiency $\eta(i,j)$ of SAL

3.4 IMPACT OF LOG-NORMAL SHADOWING

Under shadowing effects, it is likely that the monitoring areas of MoSs are not circular and could lead to anomalous situations not considered under unit model assumption. In this section, we will investigate the errors that are caused by the lack of any such information with SAL.

According to log-normal shadowing model, the power of the received signal at the MoS, P_r, is calculated from a standard path-loss model [76] as,

$$P_r = P_t - 10\alpha \log(d_{ij}) + X, \tag{3.34}$$

where P_t is the MN's transmit power, α is the path-loss coefficient, d_{ij} is the distance between MoS_i and MN_j, and $X = N(0, \sigma)$ is a normal random variable with zero mean and standard deviation σ representing the effect of shadowing on the received power. Since the localization system does not capture channel parameters, while calculating the monitoring range d, it assumes $X = N(0, 0)$. A MN at distance d_{ij} can be detected if P_r is equal to at least the detection threshold of the MoS, R_{th} (R_{th} must be at least at the receiver sensitivity level).

$$\begin{aligned} P_r &\geq R_{th} \\ P_t - 10\alpha \log(d_{ij}) + X &\geq P_t - 10\alpha \log(d) \\ -10\alpha \log(d_{ij}) + X &\geq -10\alpha \log(d). \end{aligned} \tag{3.35}$$

We assume that P_t is known by the localization system and fixed for all mobile devices; and R_{th} can be tuned so that the required monitoring range, d, can be achieved. X is unknown to localization system; therefore, the accuracy and availability of the estimation will be affected by X.

The approach in the previous sections are identical to the approach taken in this section if $\sigma = 0$. However, as discussed in previous sections, when $\sigma > 0$, the availability with SAL and CA may change. (1) If a MN is detected in a valid subarea, it can be located via SAL or CA. Availability in this case is 100% for both CA and SAL. (2) If a MN is detected in an invalid subarea, then SAL cannot estimate the location; however, CA can. Therefore, availability in this case is 0% for SAL although these regions are in the coverage area of

MoSs, and 100% for CA. (3) When the coverage is not 100% due to channel conditions or monitoring range, some fraction of MNs may not be detected at all, under which the availability is 0% for both CA and SAL.

We evaluate (via simulations) the error performance using only SAL, using CA on top of SAL (SAL+CA), and using only CA. When only SAL is used, MNs detected in valid subareas are localized using SAL, and the remaining cannot get location estimations. When SAL+CA is used, SAL is used to localize MNs detected in valid subareas, and CA is used to localize MNs detected in invalid subareas. When only CA is used, all nodes detected in valid and invalid subareas are localized using the centroid of the locations of the MoSs that hear the MNs.

In the simulations, 50,000 MNs were uniformly distributed in an area A of size 10×10. The path-loss coefficient was $\alpha = 3$ and $P_t = 10$ dBm. (In the simulations in Section IV, 10^6 nodes were used – hence some minor variations are observed between those results and the ones seen here for $\sigma = 0$.) For $\sigma > 0$, multiple simulations were run, the regions that E_{avg} achieved in these simulations fall into the regions shown in Figure 3.17; therefore, the difference in the results were insignificant.

Figures 3.18 - 3.21 show the minimum achievable average error and corresponding monitoring range, availability and overall efficiency with respect to σ for $N = 4$ and $N = 9$ MoSs in grid placement, $< N, \mathbf{G}, A, sh >$. The error reported in Figure 3.18 is the smallest mean localization error over different values of d. As expected, the error increases with σ which causes MNs to be localized erroneously. SAL or SAL+CA outperform CA significantly in accuracy. When σ is increased from 0 to 5, the increase in the minimum mean error is by $0.11L$ with 4 MoSs, and by $0.07L$ with 9 MoSs with SAL. Using SAL+CA results in a slightly higher error due to a higher number of localized MNs, and these are MNs that fall into invalid subareas and are localized with CA. Since no intersection area is available for such MoSs, the location estimates can be fairly far off. On the other hand, using only CA has a worse error performance than using SAL or SAL+CA; this justifies the effort in partitioning the network into distinct subareas and using SAL.

Figure 3.19 shows the monitoring range $\hat{d}$ which minimizes the mean localization errors for MNs in area A for different values of σ for $N = 4$ and $N = 9$. From this figure, we observe

that the best monitoring range $\hat{d}$ decreases with σ for SAL. The best value of d converges to $\sqrt{2}g/2$ which is the point when the network is barely covered by the MoSs assuming the unit disk model compared to the best value of $0.768L$ and $0.431L$ for $N = 4$ and $N = 9$ MoSs, respectively.

In Figure 3.20, we observe a significant decrease in availability with SAL. This is due to MNs detected in invalid subareas. However, SAL+CA availability stays close to 100% as CA is used for MNs detected in invalid subareas. Therefore, we can say that 40% of MNs can get location estimations with an accuracy of $0.3L$ using SAL, whereas the rest are estimated with an accuracy of $0.35L$ using CA. At high σ, CA can provide slightly better availability than SAL+CA. This is due to the increase in monitoring range with CA that is minimizing the average error. Monitoring range of SAL+CA algorithm approaches $\sqrt{2}g/2$ in order to provide the best accuracy.

According to Figure 3.21, the overall efficiency with SAL is better than with CA if $\sigma \leq 4$ with $N = 4$, and $\sigma \leq 3$ with $N = 9$, otherwise, SAL+CA outperforms CA and SAL. The results show that employing only SAL reduces the availability of localization. The reason is that with $N = 9$ MoSs, there can be many more intersection areas which are not defined as valid subareas by the unit disk model; therefore, there is a higher possibility that a MN will not be detected in a subarea. Although with $N = 9$, availability can be as low as 35% as shown in Figure 3.20, SAL+CA can provide $0.25L$ accuracy at $\sigma = 10$. Due to the same reason, the efficiency with $N = 9$ is worse than efficiency with $N = 4$ under the same channel conditions. Note that we use the same weights for localization error and the number of MoSs while calculating efficiency, i.e., $i = j = 1$.

Figures 3.22 - 3.24 show the maximum achievable availability and corresponding monitoring range and average error with respect to σ for $N = 4$ and $N = 9$. According to Figure 3.22, the maximum availability of at least 90% can be achieved with 4 MoSs when $\sigma \leq 10$, whereas with 9 MoSs the availability is less than 80% when $\sigma \geq 6$. Figure 3.23 shows that to achieve a higher availability, d increases when $N = 4$, however when $N = 9$ in order to decrease the chance to detect in invalid subareas, the monitoring range first increases and then starts to decrease.

From Figure 3.24, we observe that with $< 4, \mathbf{G}, A, sh >$, the accuracy achieved at max-

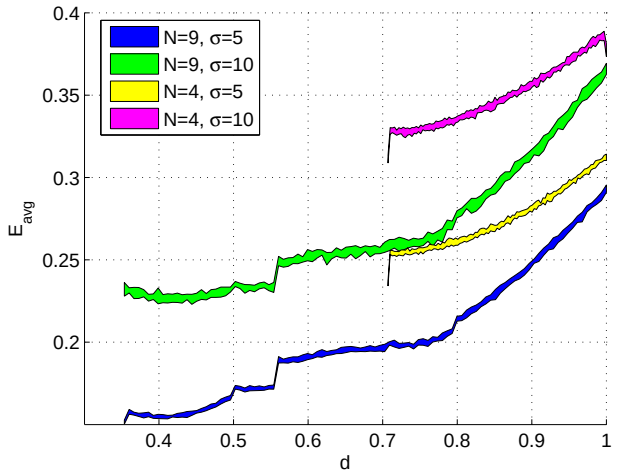

Figure 3.17: Minimum average error variance with $< 4, \mathbf{G}, A, sh >$ and $< 9, \mathbf{G}, A, sh >$

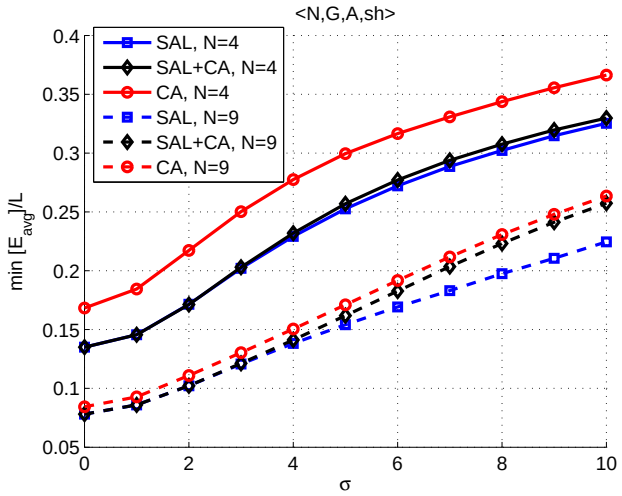

Figure 3.18: Minimum E_{avg} with $< 4, \mathbf{G}, A, sh >$ and $< 9, \mathbf{G}, A, sh >$

imum availability conditions with SAL is comparable to accuracy with CA due to increase in monitoring range. With $< 9, \mathbf{G}, A, sh >$, monitoring range of SAL does not increase as much as the monitoring range of SAL+CA; therefore, accuracy is better in spite of lower availability.

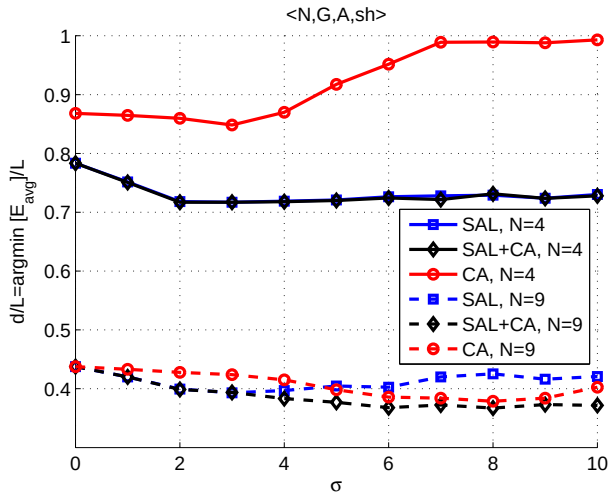

Figure 3.19: $\frac{d^*}{L} = argmin(E_{avg})$ with $< 4, \mathbf{G}, A, sh >$ and $< 9, \mathbf{G}, A, sh >$

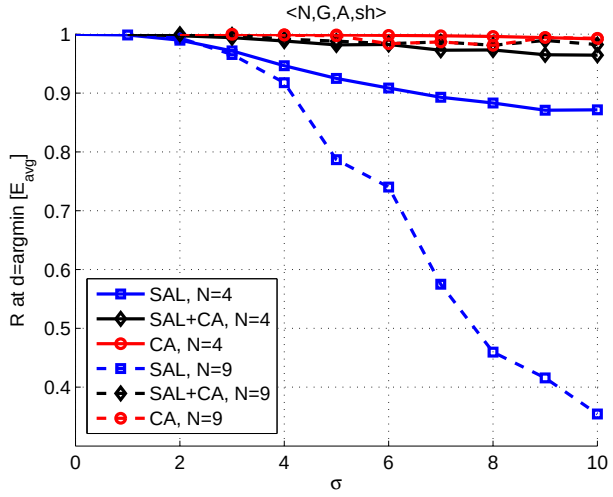

Figure 3.20: Availability at $d^* = argmin(E_{avg})$ with $< 4, \mathbf{G}, A, sh >$ and $< 9, \mathbf{G}, A, sh >$

According to Figure 3.25, the monitoring range maximizing the availability causes SAL+CA to have the best overall efficiency among the three algorithms when $\sigma \leq 4$ with $< 4, \mathbf{G}, A, sh >$ scenario, and SAL outperforms other two algorithms in the $< 9, \mathbf{G}, A, sh >$ scenario. Another observation is that 4 MoSs can provide better efficiency than 9 MoSs especially when

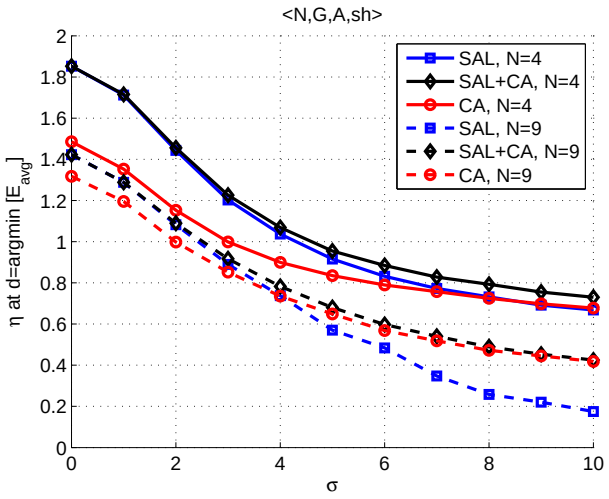

Figure 3.21: η at $d^* = argmin(E_{avg})$ with $< 4, \mathbf{G}, A, sh >$ and $< 9, \mathbf{G}, A, sh >$

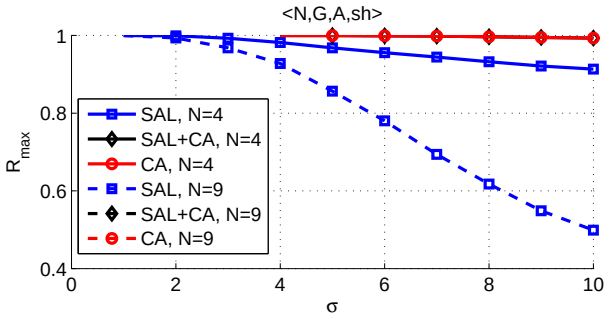

Figure 3.22: Maximum availability with $< 4, \mathbf{G}, A, sh >$ and $< 9, \mathbf{G}, A, sh >$

shadowing level is high.

Impact of Other Path-Loss Models: Although our objective is to understand the impact of X rather than being very accurate in characterizing radio propagation, more complicated/accurate path-loss models can be employed with similar results. For example, according to the Xia-Bertoni model [77], the attenuation in signal strength depends on the diffraction loss from rooftops to the street, and multiple screen diffraction past rows of buildings in addition to free space path loss. The loss is affected by mobile and base station

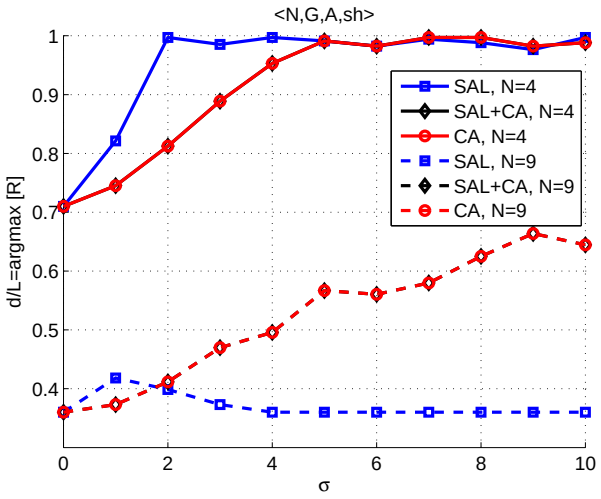

Figure 3.23: $\frac{d^*}{L} = argmax(R)$ with $< 4, \mathbf{G}, A, sh >$ and $< 9, \mathbf{G}, A, sh >$

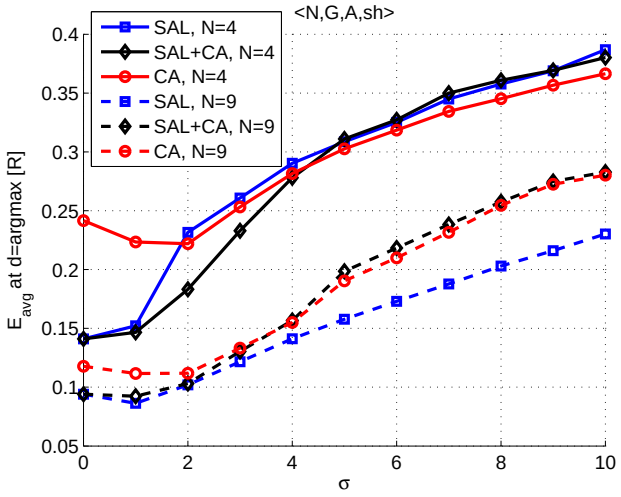

Figure 3.24: Average error at $\frac{d^*}{L} = argmax(R)$ with $< 4, \mathbf{G}, A, sh >$ and $< 9, \mathbf{G}, A, sh >$

antenna heights, height of the buildings and frequency that the signal is emitted. In [77], an estimation of loss at a distance of d is given as

$$P_L(d) = K + \beta log_{10}(f) + 10\alpha log_{10}(d).$$

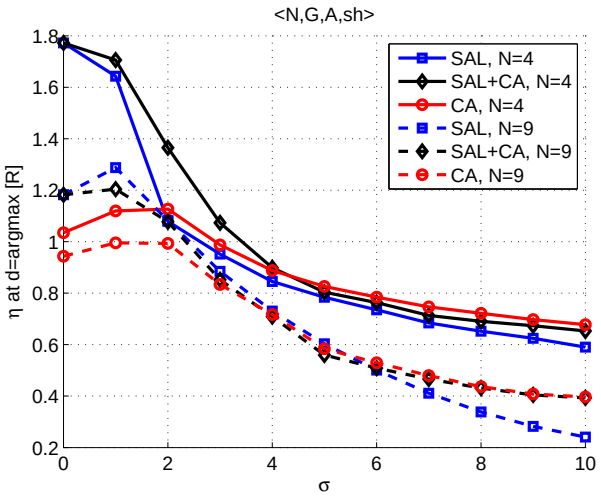

Figure 3.25: Average error at $\frac{d^*}{L} = argmax(R)$ with $< 4, \mathbf{G}, A, sh >$ and $< 9, \mathbf{G}, A, sh >$

Assuming f is fixed, and d is the monitoring range, according to this model SAL would decide if the signal is detected if $P_r \geq P_t - P_L(d)$:

$$\begin{aligned} P_t - P_L(d_{ij}) + X &\geq R_{th} \\ -10\alpha log_{10}(d_{ij}) + X &\geq -10\alpha log_{10}(d), \end{aligned} \tag{3.36}$$

where we can clearly see that the criteria for detection of MNs stays the same as the simple model with log normal shadowing that we previously used.

3.4.1 Effect of Wall Attenuation Factor

The use of Wall Attenuation Factor in localization was proposed by Bahl et. al in [3]. According to the model used there, each time a signal passes through a wall, the received power decreases by WAF in dB. When we also consider the shadowing, the received power at a monitoring station is:

$$P_r = P_t - 10\alpha \log(d_{ij}) + X - n\ WAF \tag{3.37}$$

where n is the number of walls that the signal is passing through, and X is log-normal shadowing with mean 0 and standard deviation σ.

We assume that the localization system *knows* the locations of walls, the attenuation in dB per wall, and the locations of monitoring stations. Then, monitoring stations may expect the signals to arrive passing through an average number of walls. Then, the localization algorithm uses the following formula to decide that a mobile is detected if the received signal power is at least R_{th}:

$$P_r \geq R_{th} = P_t - 10\alpha \log(d) - n_{avg}\, WAF \tag{3.38}$$

where n_{avg} is the average number of walls that the signal is assumed to be passing through, and d is planned monitoring range.

In order to show the effect of wall attenuation factor, we worked with two different floor plans (FP 1 and FP 2) shown in Figures 3.26 (a) and (b). In each case, 4 MoSs are deployed at the corners. Black lines represent the border of localization service area, and green lines represent the walls. In the first floor plan, there is a hallway with three rooms on each side, and in the second floor plan, there are rooms on all sides, and there is another room in the middle. We assume $WAF = 3.1$dB which was empirically calculated for WiFi at 2.4GHz in [3], and similar attenuation is seen over windows and doors.

Figures 3.27 and 3.28 show the minimum achievable average error and corresponding monitoring range with respect to σ. Comparing the accuracy performance in Figures 3.27, and Figure 3.18, we observe that the accuracy in FP1 was not affected, whereas the accuracy in FP2 gets worse at low shadowing levels. From Figure 3.28, we see that d is decreasing with increasing σ and approaches to $\sqrt{2}g/2$. The availability when the minimum error monitoring range is employed (in Figure 3.29) is close to 100% if $\sigma \leq 2$, and it is decreasing gradually down to 65% at $\sigma = 10$. However, with SAL+CA a larger monitoring range is used and availability can be increased up to 95% at $\sigma = 10$. It is important to note that although it is observed that at $\sigma = 10$ better accuracy is achieved with SAL when compared to Figure 3.18, this is due to lower availability of location estimation as seen from Figure 3.29.

We can observe from Figure 3.30 that the overall efficiency of SAL+CA is better than the others due to availability and accuracy performance. When compared to $< 4, \mathbf{G}, A, sh >$

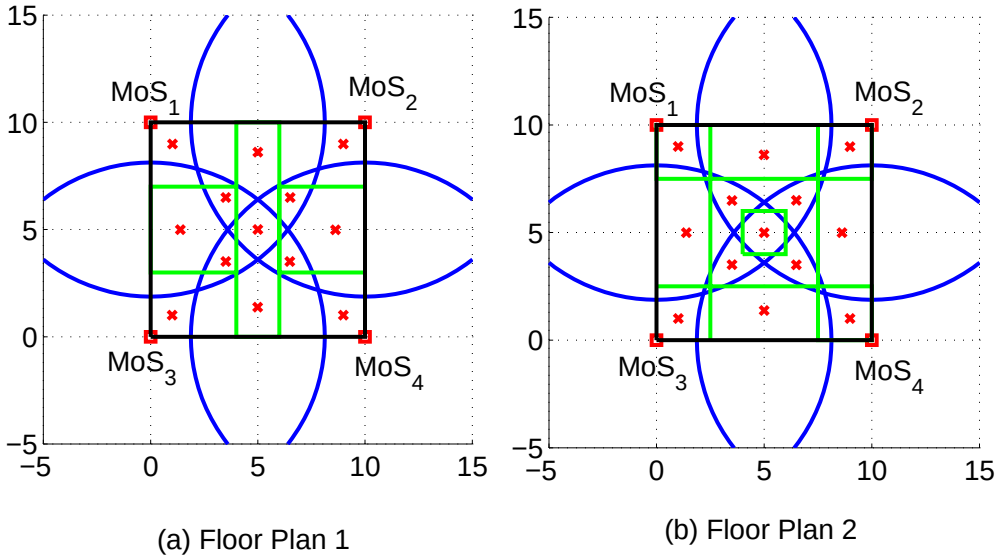

Figure 3.26: (a) Floor Plan 1 (FP 1) (b) Floor Plan 2 (FP 2)

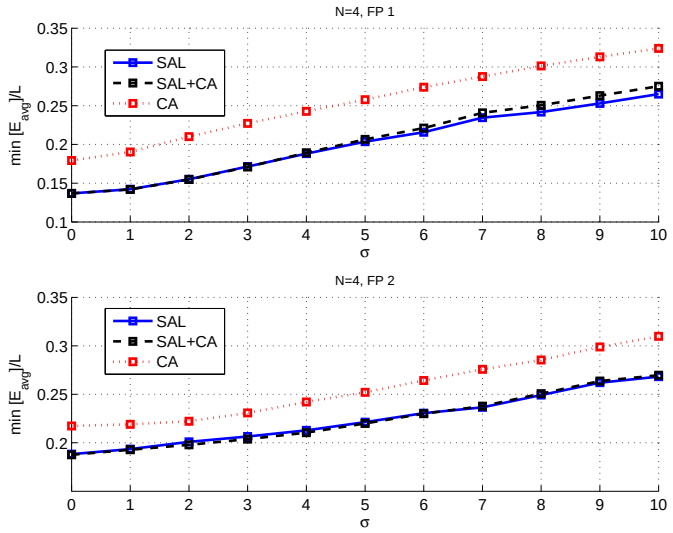

Figure 3.27: Minimum average error with $< 4, \mathbf{G}, A, waf >$ on FP1 and FP2

we do not see a significant reduction in efficiency for FP 1, but we see worse efficiency in

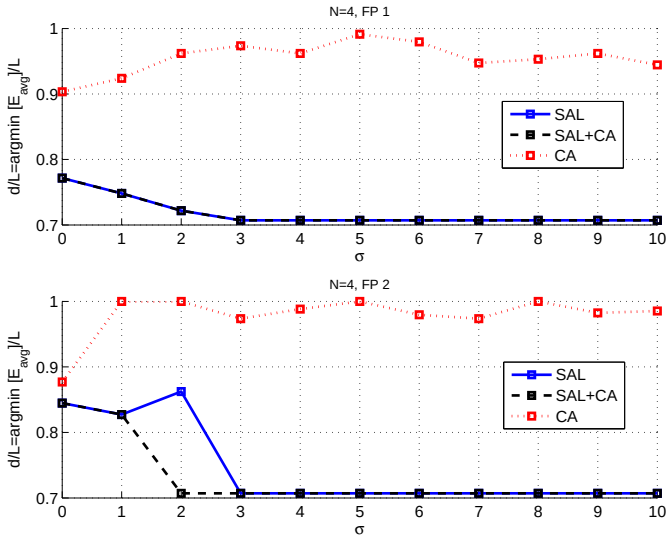

Figure 3.28: $\frac{d^*}{L} = argmin(E_{avg})$ with $< 4, \mathbf{G}, A, waf >$ on FP1 and FP2

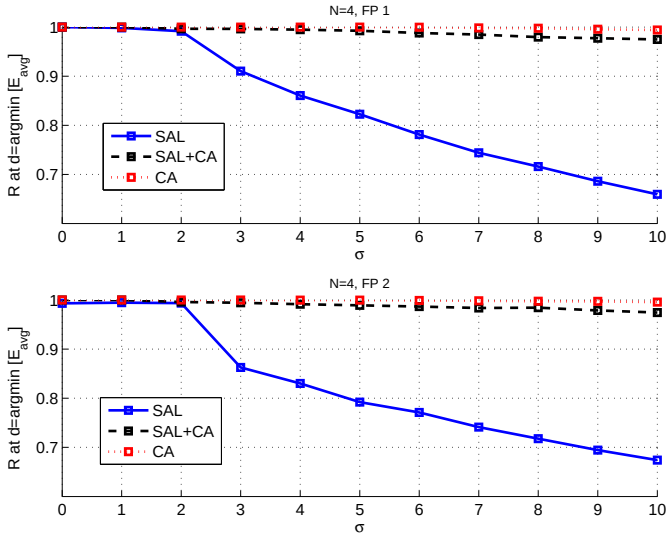

Figure 3.29: Availability at d^* with $< 4, \mathbf{G}, A, waf >$ on FP1 and FP2

FP 2. However, the values are fairly close indicating that it may be sufficient to consider situations without walls especially when there are few walls impacting the subareas.

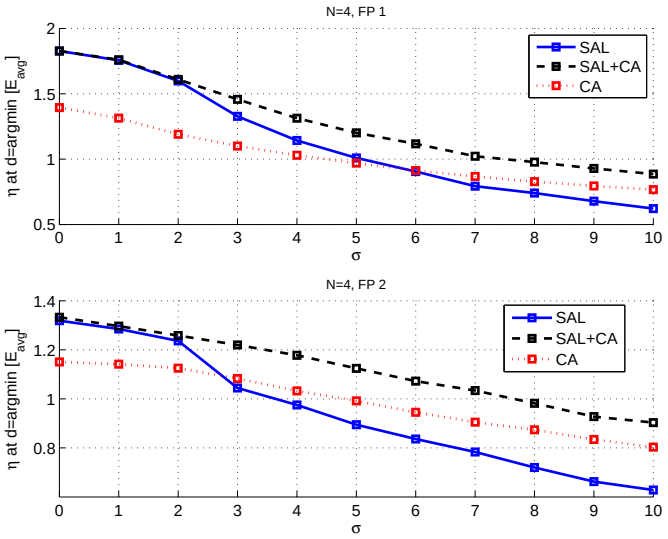

Figure 3.30: Efficiency at d^* with $< 4, \mathbf{G}, A, waf >$ on FP1 and FP2

3.4.2 Sensitivity to Multiple Detection Thresholds

Utilization of multiple thresholds for the RSS compared to a single threshold (RSS detected or not) may increase the resolution of sub-areas, and improve localization accuracy. This is similar to increasing the granularity in the location fingerprint (instead of a 0 and 1). In this section, we are investigating the effect of using multiple RSS thresholds on accuracy and availability. When r RSS thresholds are used, the MNs are detected in the r^{th} ring if:

$$R_{th}(r) \leq P_r < R_{th}(r+1), \ (R_{th}(0) = 0, \ r \geq 1)$$

A visualization of subareas with three RSS thresholds in scenario $< 4, \mathbf{G}, A >$ is shown in Figure 3.31. Blue, green and purple arcs represent the RSS threshold limits according to the unit disk model. We assumed the same radius separation for each ring.

Figures 3.32 and 3.33 show the results of using single and multiple RSS thresholds. From Figure 3.32, it seems that using multiple RSS thresholds can decrease the localization error by 33% - 50% depending on channel conditions. Accuracy with two RSS_{th}s at $\sigma = 0$ resembles the accuracy with $< 9, \mathbf{G}, A, sh >$ with a single RSS_{th}. From Figure 3.33, we see

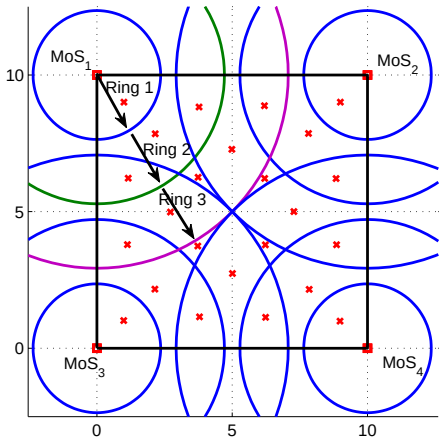

Figure 3.31: Visualization of subareas using three RSS thresholds with $< 4, \mathbf{G}, A >$

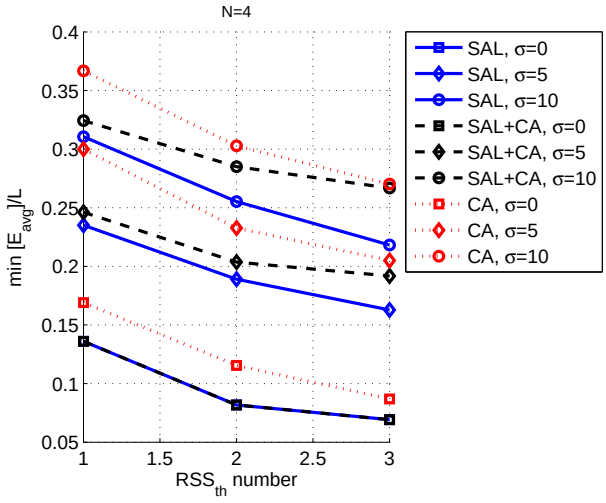

Figure 3.32: Minimum E_{avg} using multiple RSS thresholds with $< 4, \mathbf{G}, A, sh >$

that the monitoring range providing the smallest localization error increases with number of RSS thresholds.

Under ideal conditions, we expect better accuracy, as using n RSS thresholds with N MoSs can provide the subareas that can be created by nN MoSs using a single RSS thresh-

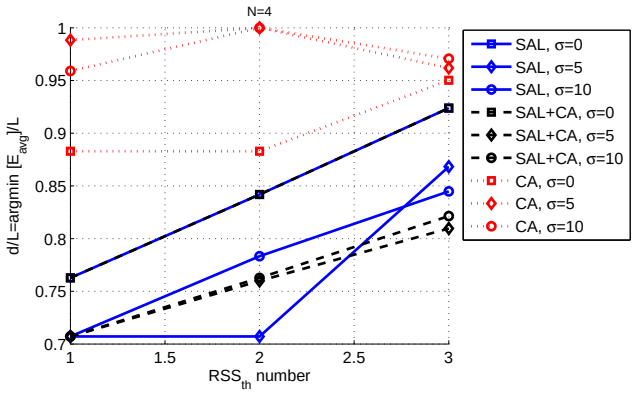

Figure 3.33: $\frac{d^*}{L} = argmin(E_{avg})$ using multiple RSS thresholds with $< 4, \mathbf{G}, A, sh >$

old. As nN MoSs may create more invalid subareas than with N MoSs under shadowing, using multiple RSS thresholds would also cause many invalid subareas leading to smaller availability. Simulation results shown in Figure 3.34 are in line with the intuition. The availability and accuracy performance reflects in the best efficiency that is seen for SAL+CA among the three algorithms.

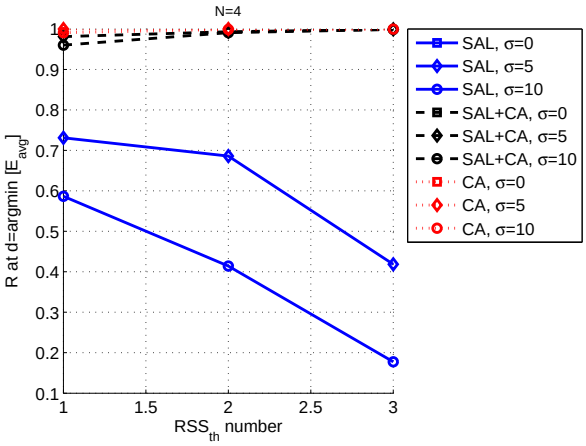

Figure 3.34: Availability at $\frac{d^*}{L}$ using multiple RSS_{th}s with $< 4, \mathbf{G}, A, sh >$

We note here that using multiple RSS thresholds comes at a calibration cost of correctly determining the ranges at these thresholds during deployment and tuning the thresholds -

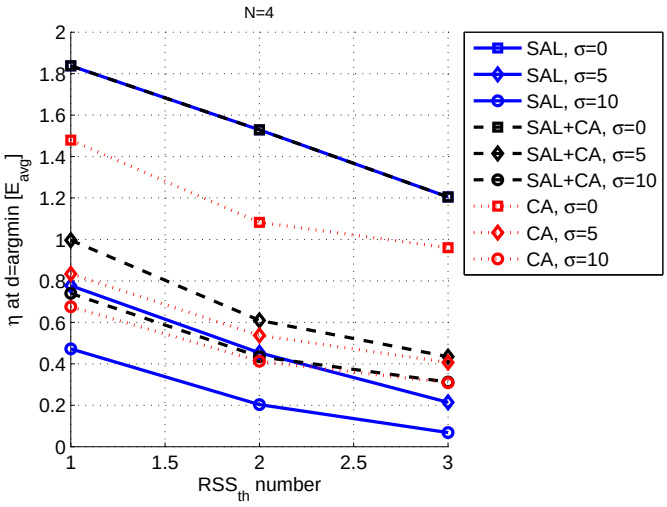

Figure 3.35: Efficiency at $\frac{d^*}{L}$ using multiple RSS_{th}s with $< 4, \mathbf{G}, A, sh >$

something that has not been considered in this work. Thus we have not assigned a cost for this effort in the above evaluations.

3.4.3 Incremental Addition of Monitoring Stations

Access points are widely deployed and already used for wireless services. Therefore, we investigate a scenario where there are already available access points in the localization area, to determine how the performance may improve when additional access points are deployed. Thus, in this section, we are assuming 4 MoSs were already deployed at the corners of the localization area, and other 5 MoSs are incrementally deployed one by one in the order shown in Figure 3.7. When all MoSs are finally deployed, this scenario will be equivalent to the "9 MoSs on a grid scenario". We assume the allowed monitoring ranges with different number of MoSs will ensure 100% availability under ideal channel conditions; therefore, for $N = 4$ MoSs, $d \geq \sqrt{2}L/2$; for $N = 5, 6, 7, 8$, $d \geq L/2$; and for $N = 9$, $d \geq \sqrt{2}L/4$.

Figure 3.36 shows that the accuracy gets better as more MoSs are used in localization. The MoSs added to the system are spread in the localization area; therefore, every time a MoSs is added, an improvement in accuracy is observed. The best monitoring ranges

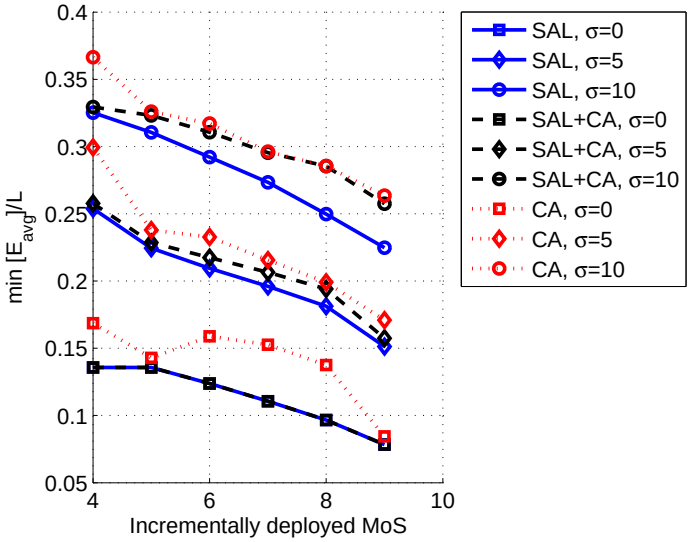

Figure 3.36: Minimum average error with $< 4, \mathbf{G}, A, sh >$

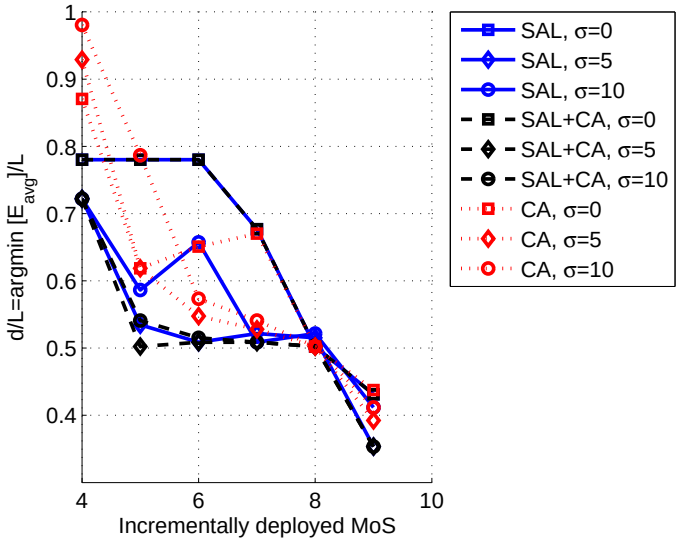

Figure 3.37: $\frac{d^*}{L} = argmin(E_{avg})$ with $< 4, \mathbf{G}, A, sh >$

at each increment are shown in Figure 3.37. As also discussed before, the availability of location estimation is reduced under shadowing with increasing number of MoSs as shown

in Figure 3.38. Figure 3.39 shows the overall efficiency improves only under ideal channel conditions as N increases; however, when $\sigma > 0$, using smaller number of MoSs is more efficient.

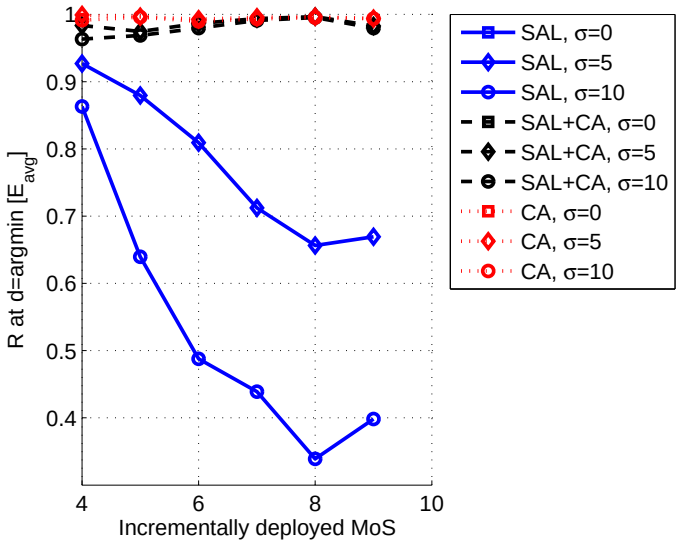

Figure 3.38: Availability at $d^* = argmin(E_{avg})$ with $< 4, \mathbf{G}, A, sh >$

3.5 PERTURBATION FROM GRID PLACEMENT

In our previous analyses, we assumed a grid placement for monitoring stations for the sake of simplicity in focusing on specific aspects of the problem. However, due to the environment, the furniture, walls or other reasons, placing the monitoring stations so that they will form a grid may not be possible. Since the performance of our localization approach depends on the joint monitoring areas of MoSs, we will investigate the effect of the diversion from grid placement on the accuracy of localization. We denote this scenario as $< N, \tilde{G}, A >$.

For this analysis, we use $N = 4$ MoSs grid placement as a baseline, and gradually increase the perturbation from grid placement. Perturbation is applied by deploying each MoS uniformly randomly in a square area of size $2pL \times 2pL$, of which the center is the actual

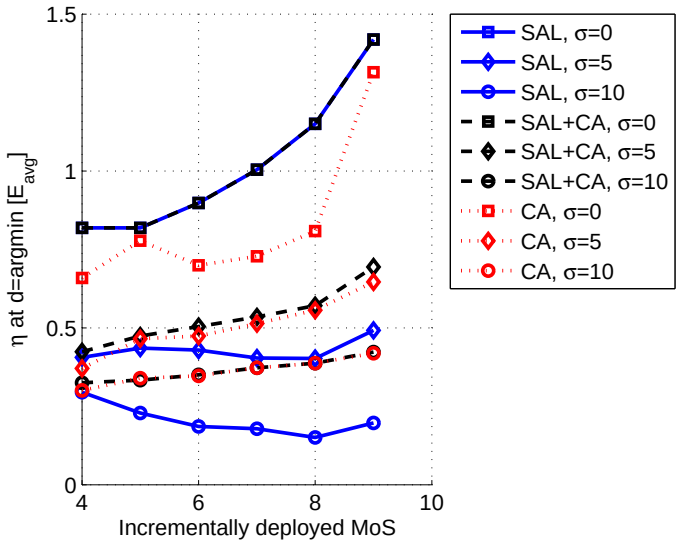

Figure 3.39: η at $d^* = argmin(E_{avg})$ with $< 4, \mathbf{G}, A, sh >$

grid point, and p is the perturbation fraction. An instance of perturbed grid with $p = 0.3$ is shown in Figure 3.40, where the modified joint areas and potential estimated locations (blue crosses) can be clearly seen. We note that the target area for the localization service is inside the black square; although the monitoring stations (in this placement) are outside this area, we are still interested in the localization performance inside the black square. Perturbation can result in imbalance among the subarea sizes, as well as smaller number of subareas (although it is not the case in Figure 3.40).

Figures 3.41 - 3.42 show the minimum achievable average error, monitoring range required to achieve this level of accuracy and availability factor at this range when grid placement of 4 MoSs are perturbed by $p = 0,\ 0.1L,\ 0.2L,\ 0.3L,\ 0.4L,\ 0.5L$. Every point in the plots is the average of 20 randomly generated perturbation scenarios. In order to single out the effects of perturbation, we assumed log-normal shadowing without WAF.

From Figure 3.41, when $\sigma = 0$, with SAL the increase in minimum average error is 21% with $p = 0.5L$ when compared to grid placement ($p = 0$), whereas with CA, the accuracy gets worse by 75%. When $\sigma = 10$, the accuracy gets worse by 13% with SAL, and 14% with

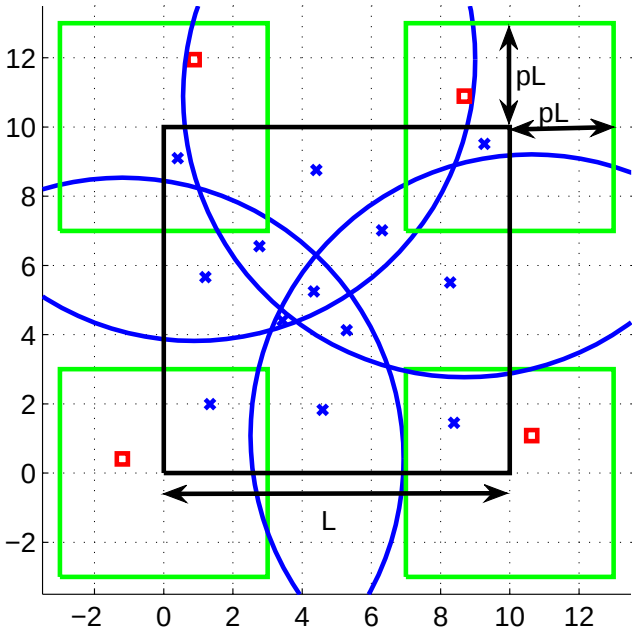

Figure 3.40: Perturbation on grid, $N = 4$

CA. For example, the accuracy gets worse by $0.03L$ when the grid is perturbed by $0.3L$ at $\sigma = 5$. When $L = 10$, this corresponds to 30cm worse accuracy, when we are allowed to change the location of monitoring stations in a 3m by 3m area which is as big as a regular office room. Under different shadowing conditions, the effect of perturbation will not change significantly.

In Figure 3.41, we observe that in order to provide this level of accuracy, monitoring range shows an increasing trend while perturbation factor increases. In spite of insignificant effect on accuracy, in Figure 3.42 SAL availability is shown to be reducing down to approximately 65% with perturbation under shadowing effect. According to the overall efficiency, using only SAL+CA under mild or severe shadowing is a good design choice in order to provide the accuracy shown in Figure 3.43.

Figures 3.44-3.46 show the maximum achievable availability, the monitoring range that provides this level of availability, and average localization error at this range under different

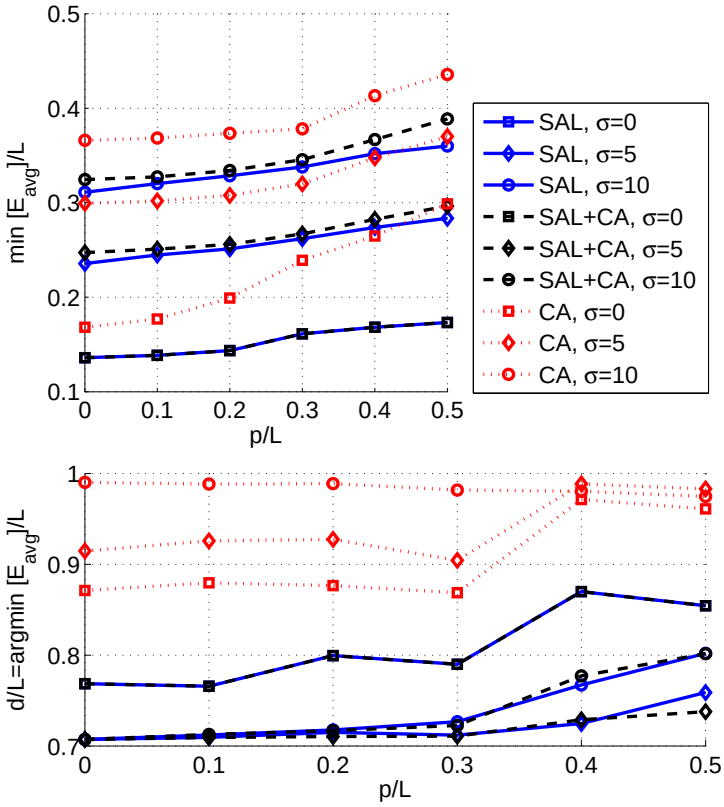

Figure 3.41: Minimum average error with $< 4, \tilde{G}, A, sh >$

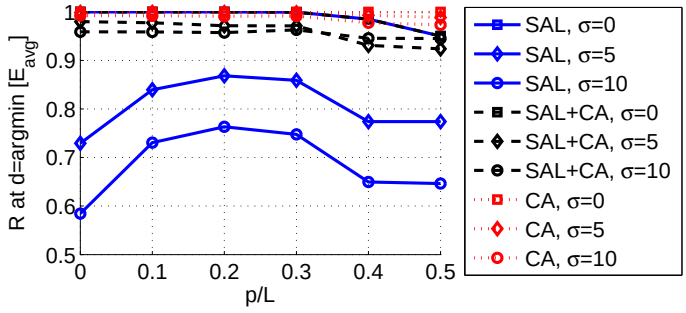

Figure 3.42: Availability at $d^* = argmin(E_{avg})$ with $< 4, \tilde{G}, A, sh >$

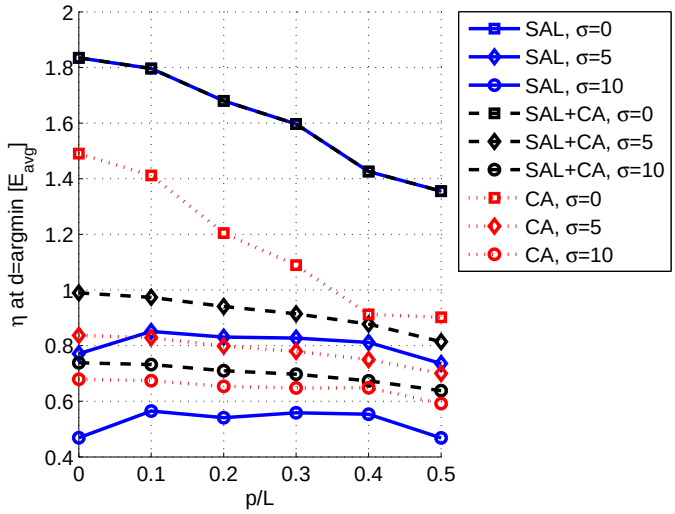

Figure 3.43: Efficiency at $d^* = argmin(E_{avg})$ with $< 4, \tilde{G}, A, sh >$

shadowing conditions, and with different perturbation fractions. Using SAL, R_{max} is around 92% with grid placement when $\sigma = 10$ and it decreases by approximately 10% when $0.5L$ perturbation on grid is allowed. CA availability, on the other hand, is slightly affected by perturbation.

Figure 3.45 shows the corresponding monitoring ranges. At higher shadowing levels, monitoring range for SAL decreases in order to minimize the invalid subarea detection probability. According to Figure 3.46, accuracy of SAL at corresponding monitoring ranges is not affected by perturbation significantly as it stays under $0.39L$ at high shadowing levels; otherwise, it increases with increasing perturbation factor. As a result, the overall efficiency suggests that a deployment employ SAL when there is no shadowing, and SAL+CA with mild shadowing, and CA with high shadowing as shown in Figure 3.47.

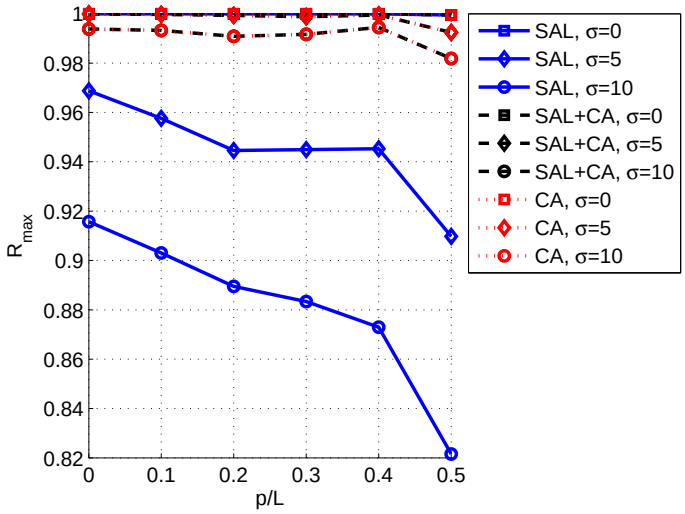

Figure 3.44: Maximum availability with $< 4, \tilde{G}, A, sh >$

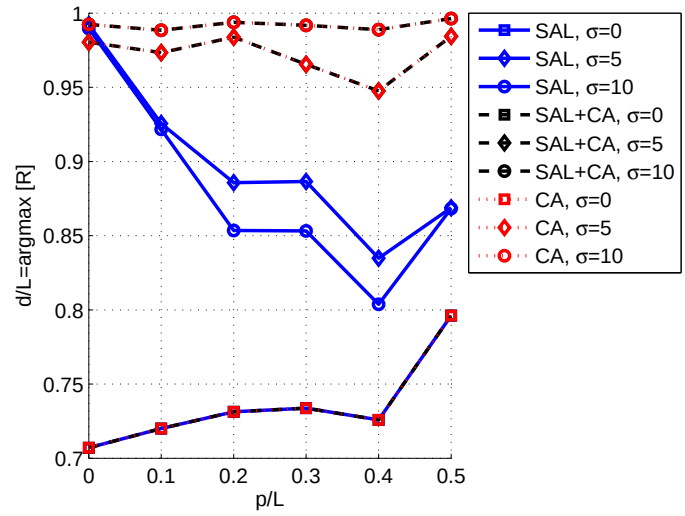

Figure 3.45: $\frac{d^*}{L} = argmax(R)$ with $< 4, \tilde{G}, A, sh >$

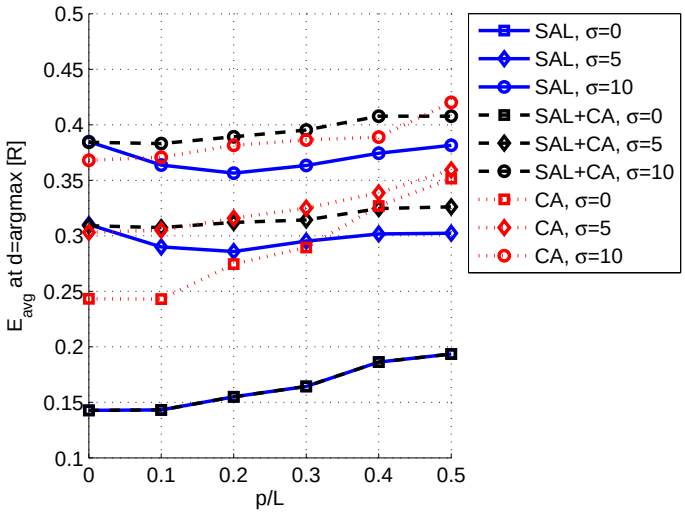

Figure 3.46: Average error at $\frac{d^*}{L} = argmax(R)$ with $< 4, \tilde{G}, A, sh >$

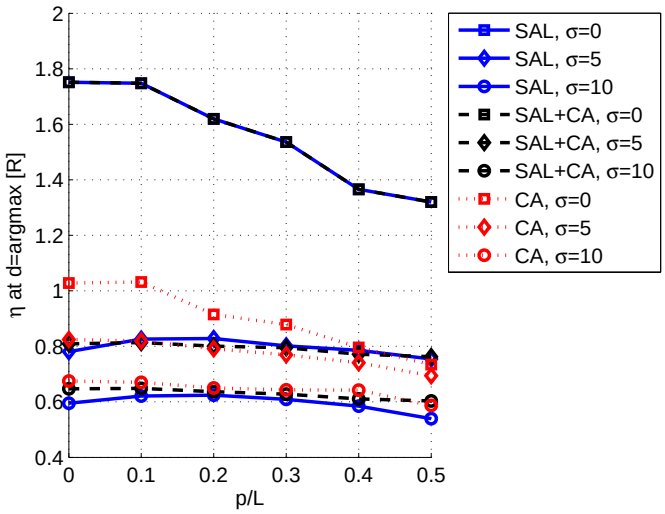

Figure 3.47: Efficiency at $\frac{d^*}{L} = argmax(R)$ with $< 4, \tilde{G}, A, sh >$

3.6 PERFORMANCE UNDER RANDOM PLACEMENT

In the previous sections, we investigated the performance of SAL when MoSs are placed on a virtual grid and how the performance is affected by the perturbation of this grid.

Both these placements require some effort which will in turn reflect as higher installation cost. Randomly placing MoSs requires much less effort as we do not have to align the locations of MoSs so that they will form a grid. However, sub-area shapes, sizes and numbers will be different and unpredictable when compared to grid placement. This can affect the localization performance negatively. Therefore, in this section, we want to determine the trade off between the accuracy and practical deployment through simulations with randomly placed MoSs. In order to see the effect of random placement, instead of randomly placing N MoSs into the whole area, it is better to divide the whole area into smaller areas, and then place MoSs randomly inside those areas. This will provide a more homogeneous random placement especially when N is small. We denote this scenario as $< N, \mathbf{R}, A >$. An instance of such placement of 4 MoSs in 4 different square areas of size $L/2 \times L/2$ in area of size $L \times L$ is shown in Figure 3.48. Dashed lines virtually divide the whole area into 4 square areas. From this figure, it is clearly seen that the subarea sizes, shapes and numbers change significantly with random placement when compared to grid placement shown in previous sections. If centroid algorithm is used with this type of placement, the potential estimated locations will always be inside the region that is defined by the line connecting the locations of MoSs. However, with SAL, the potential estimated locations are more spread into the whole area as shown with red crosses in Figure 3.48, and they do not have to be inside this specific region. Intuitively thus, SAL can provide better accuracy.

In the following, we show the effect of random placement by simulations of two localization scenarios using 4 and 8 MoSs, $< 4, \mathbf{R}, A, sh >$ and $< 8, \mathbf{R}, A, sh >$ assuming log-normal shadowing. The results shown in this section are the average of 30 such instances of random placement of MoSs.

Figure 3.49 shows the best accuracy that can be achieved by tuning the monitoring ranges under different shadowing conditions. When $< 4, \mathbf{R}, A, sh >$ is compared to $< 4, \mathbf{G}, A, sh >$, the accuracy gets worse by only $0.01L$ under ideal channel conditions. We also observe that accuracy performance with $< 8, \mathbf{R}, A, sh >$ is comparable to performance with $< 9, \mathbf{G}, A, sh >$ as shown in Section 3.4. While σ is increasing, the localization error increases and approaches the error with CA. In the best random placement case, 4 MoSs are placed at the center of each square. With this placement, the monitoring range must be at least

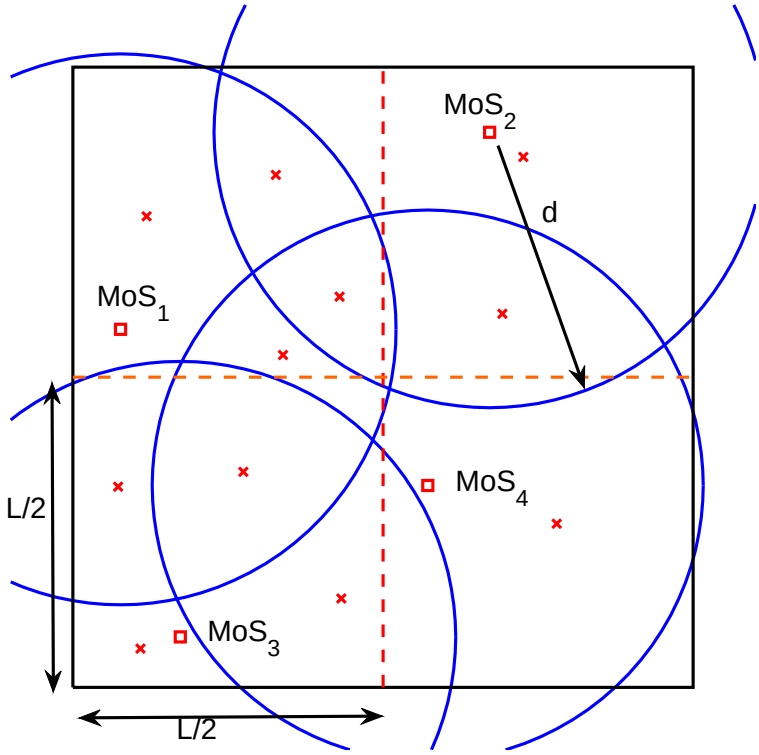

Figure 3.48: Random placement of $N = 4$ MoSs in $L \times L$

$\sqrt{2}L/4$ for the coverage to be 100%. Therefore, we limited the minimum monitoring range to this value. And this is the reason why the monitoring range providing the best accuracy is approaching $\sqrt{2}L/4$ under higher σ as shown in Figure 3.50. With $< 8, \mathbf{R}, A, sh >$, the monitoring range providing the best accuracy first decreases and then increases for increasing σ, this behavior is the same as the behavior observed with $< 9, \mathbf{G}, A, sh >$.

In the worst case, all MoSs can be placed at the corners of the square shaped area, and $d \geq \sqrt{2}L/2$ for full coverage. Therefore, when $d < \sqrt{2}L/2$, the coverage of MoSs may not be 100% when MoSs are randomly placed. For example in Figure 3.48, the lower right corner is not covered by any of the MoSs although $d > \sqrt{2}L/4$. This is reflected on the location

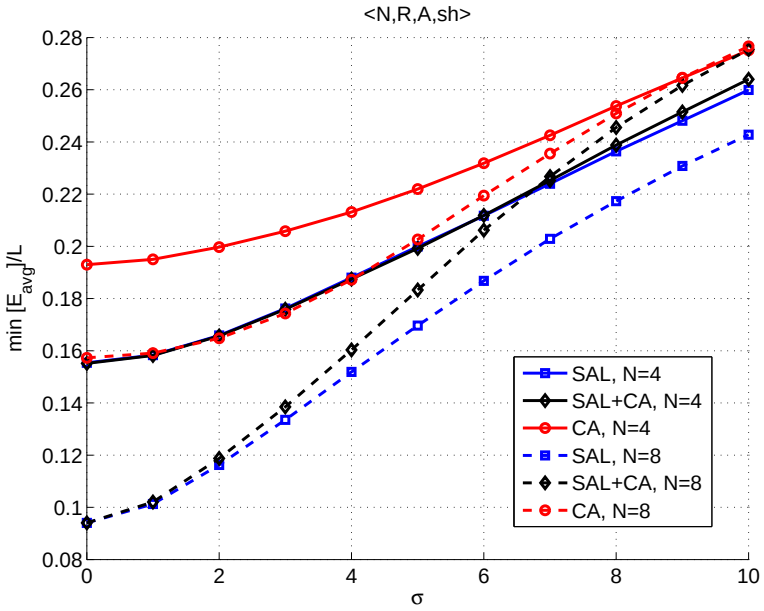

Figure 3.49: Minimum E_{avg} with $< 4, \mathbf{R}, A, sh >$ and $< 8, \mathbf{R}, A, sh >$

estimation availability performance as shown in Figure 3.51. We observe that even under ideal channel conditions, availability is around 93% with 4 MoSs and 98% with 8 MoSs, and it is decreasing gradually down to 70% and 40% with increasing shadowing effects. This behavior is also observed with $< 4, \mathbf{G}, A, sh >$ and $< 9, \mathbf{G}, A, sh >$; therefore, availability at best accuracy settings are not affected by random placement of MoSs.

Figures 3.52 shows the maximum availability that can be achieved under different channel conditions and monitoring range required. SAL+CA and CA have the same availability and same best monitoring range providing this value. For lower σ, both SAL and CA have good availability performance. However, as σ increases, the availability with SAL starts to decrease. In order to keep availability at the maximum level, the monitoring range increases for small σ values (see Figure 3.53), and starts to decrease under higher σ in order to decrease the number of MNs detected in invalid subareas. This behavior is reflected in higher localization errors as shown in Figure 3.54.

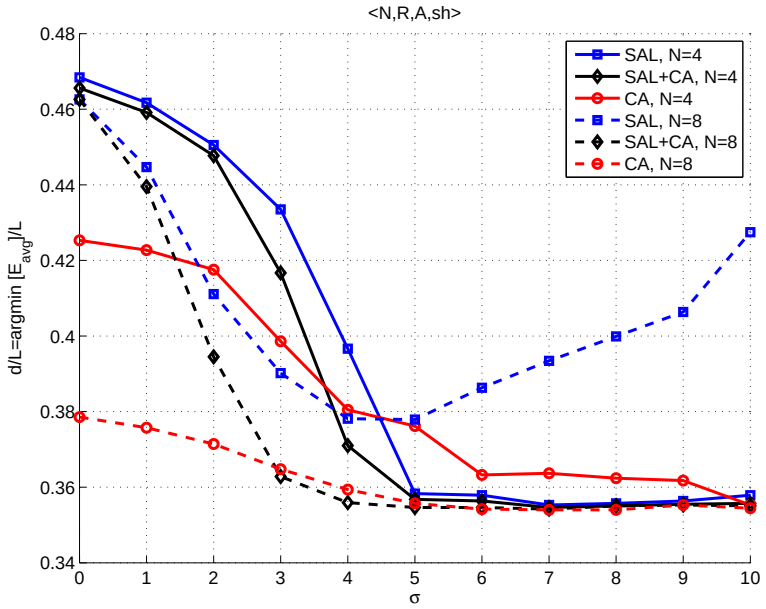

Figure 3.50: $d^* = argmin(E_{avg})$ with $< 4, \mathbf{R}, A, sh >$ and $< 8, \mathbf{R}, A, sh >$

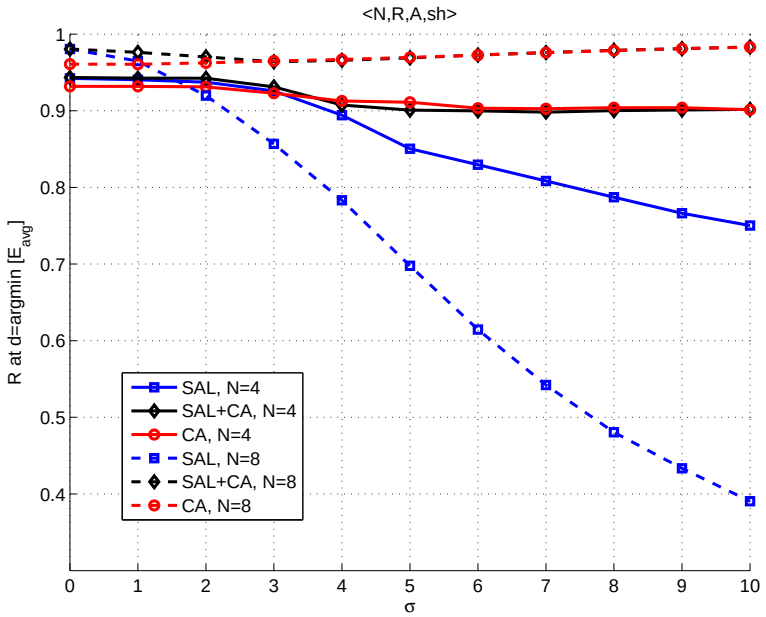

Figure 3.51: Availability at $d^* = argmin(E_{avg})$ with $< 4, \mathbf{R}, A, sh >$

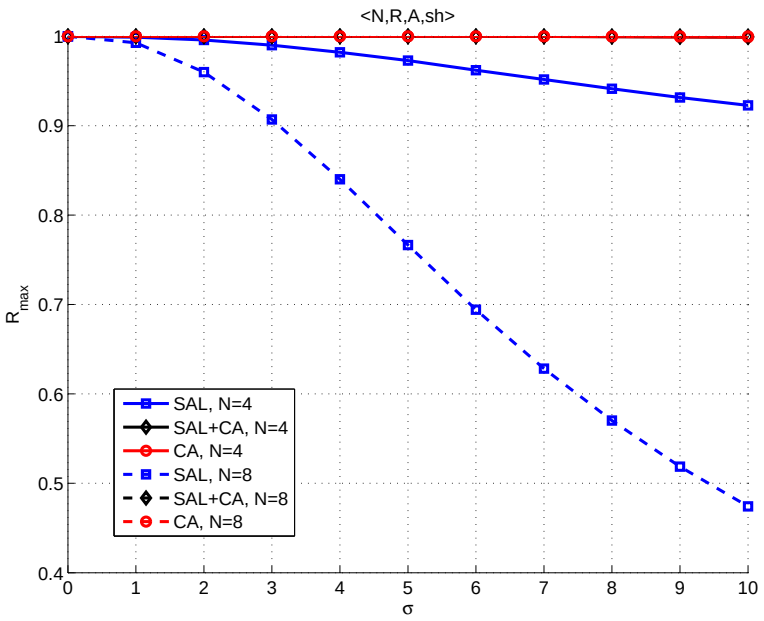

Figure 3.52: Maximum R with $< 4, \mathbf{R}, A, sh >$

3.7 CONCLUSIONS

In this chapter, we propose Sub-Area Localization which is a 2-dimensional proximity-based technique. Analytical and simulation results are used to determine the optimal coverage (range) for best localization accuracy with SAL given a localization scenario with a number of monitoring stations covering the network area. Our results indicate that there is an optimal range and as we increase the number of MoSs, the localization error and monitoring range reduce, following an exponential fit. From a practical standpoint, the clear-channel-access (CCA) threshold in WiFi can be used to tune the monitoring ranges of MoSs to create the best sub-areas for accurate localization.

We compare SAL to fine grained algorithms that has been cited in Section 2 in terms of localization error normalized with localization area in Figure 3.55. According to this, SAL as a proximity based algorithm has worse accuracy than fine grained algorithms due to its coarse grained nature. However, we show that under shadowing with standard deviation of

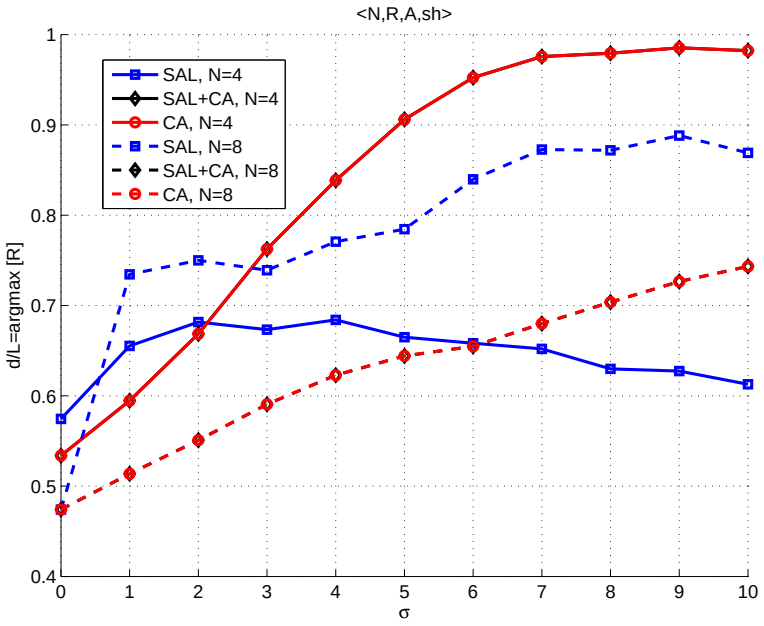

Figure 3.53: $d^* = argmax(R)$ with $< 4, \mathbf{R}, A, sh >$

3dB SAL with 9 MoSs can have a similar accuracy as WS algorithm which is using wireless sniffers to gather information regarding the current channel parameters. Nevertheless, it is important to emphasize that SAL does not use any communication between monitoring stations to capture channel parameters; therefore, it is more likely that online-calibration based algorithms have less variance in terms of accuracy.

Tables 3.6, 3.7 and 3.8 show the minimum localization and best monitoring range of SAL, CA, and SAL+CA algorithms in various localization scenarios and channel conditions. From these tables, we can conclude that,

- SAL has a significantly better accuracy than with localization that uses the centroid of the locations of MoSs under different channel conditions and localization scenarios.

- The best monitoring range with grid placement scenario can also provide the best accuracy conditions for localization in scenarios when grid placement of MoSs is not possible.

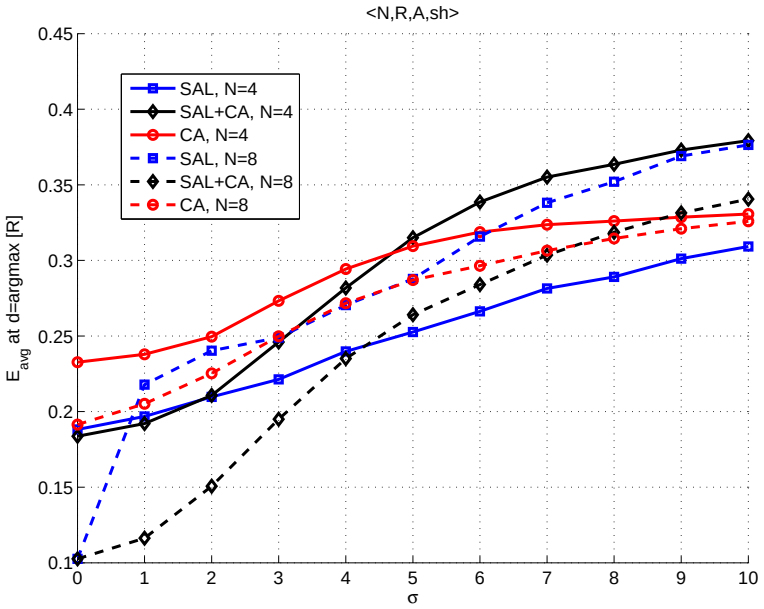

Figure 3.54: E_{avg} at $d^* = argmax(R)$ with $< 4, \mathbf{R}, A, sh >$

- The centroid localization can help increase the availability of localization when used with SAL in SAL+CA. From the results, SAL+CA seems to have a slightly worse accuracy than SAL. The reason is the availability of SAL+CA is larger than availability of SAL, which reflects in a higher average error.

- When wall attenuation factor is considered in addition to log normal shadowing, SAL+CA has similar availability to CA availability, and similar accuracy to SAL accuracy, making this algorithm the one that has the best localization efficiency.

- When wall attenuation factor is considered, different accuracy levels are achieved with different floor plans. The difference becomes negligible as shadowing conditions become significant and thus may be ignored, but SAL can once again improve the accuracy compared to CA.

- Utilizing multiple detection thresholds can help significantly improve accuracy of MNs

Table 3.6: SAL accuracy with $N = 4$ MoSs in various localization scenarios

Scenario		$\sigma = 0$	$\sigma = 5$	$\sigma = 10$
$< 4, \mathbf{G}, A, sh >$	E^*_{avg}/L	0.135	0.26	0.33
	d^*/L	0.768	0.707	0.707
$< 4, \mathbf{G}, A, waf >$	E^*_{avg}/L	0.14	0.2	0.264
FP 1	d^*/L	0.77	0.707	0.707
$< 4, \mathbf{G}, A, waf >$	E^*_{avg}/L	0.19	0.22	0.27
FP 2	d^*/L	0.84	0.707	0.707
$< 4, \tilde{\mathbf{G}}, A, sh >$	E^*_{avg}/L	0.16	0.26	0.34
$p = 0.3L$	d^*/L	0.78	0.71	0.72
$< 4, \mathbf{G}, A, sh >$	E^*_{avg}/L	0.08	0.16	0.23
$\lvert RSS_{th}\rvert = 3$	d^*/L	0.92	0.87	0.85
$< 4, \mathbf{R}, A, sh >$	E^*_{avg}/L	0.15	0.2	0.26
	d^*/L	0.468	0.358	0.357

that can be localized. A drawback of this method appears under severe shadowing as it affects the availability negatively. Therefore, this design is more convenient for localization areas with close to ideal channel conditions (e.g., open spaces). Further there is a cost in determining the ranges for thresholds in real environments.

- Random placement of MoSs may affect the availability negatively for all three algorithms. However, SAL+CA has a similar availability to availability achieved with CA, and also similar accuracy to SAL's accuracy.

Table 3.7: SAL+CA accuracy with $N = 4$ MoSs in various localization scenarios

Scenario		$\sigma = 0$	$\sigma = 5$	$\sigma = 10$
$< 4, \mathbf{G}, A, sh >$	E^*_{avg}/L	0.135	0.26	0.33
	d^*/L	0.768	0.707	0.707
$< 4, \mathbf{G}, A, waf >$	E^*_{avg}/L	0.136	0.2	0.275
FP 1	d^*/L	0.77	0.707	0.707
$< 4, \mathbf{G}, A, waf >$	E^*_{avg}/L	0.19	0.22	0.27
FP 2	d^*/L	0.84	0.707	0.707
$< 4, \tilde{\mathbf{G}}, A, sh >$	E^*_{avg}/L	0.16	0.26	0.345
$p = 0.3L$	d^*/L	0.78	0.71	0.72
$< 4, \mathbf{G}, A, sh >$	E^*_{avg}/L	0.08	0.19	0.27
$\lvert RSS_{th}\rvert = 3$	d^*/L	0.92	0.81	0.82
$< 4, \mathbf{R}, A, sh >$	E^*_{avg}/L	0.155	0.2	0.264
	d^*/L	0.465	0.356	0.355

- Finally, the results of how SAL and SAL+CA performs under grid placement, ideal channel conditions, and under shadowing (and with monitoring ranges that barely cover an area) seem to be sufficient to obtain reasonably good estimates of localization errors, availability, and efficiency under more complicated scenarios.

Table 3.8: CA accuracy with $N = 4$ MoSs in various localization scenarios

Scenario		$\sigma = 0$	$\sigma = 5$	$\sigma = 10$
$< 4, \mathbf{G}, A, sh >$	E^*_{avg}/L	0.168	0.3	0.366
	d^*/L	0.87	0.92	0.99
$< 4, \mathbf{G}, A, waf >$	E^*_{avg}/L	0.18	0.257	0.324
FP 1	d^*/L	0.9	0.99	0.94
$< 4, \mathbf{G}, A, waf >$	E^*_{avg}/L	0.22	0.25	0.31
FP 2	d^*/L	0.87	1	0.98
$< 4, \tilde{\mathbf{G}}, A, sh >$	E^*_{avg}/L	0.24	0.32	0.38
$p = 0.3L$	d^*/L	0.86	0.9	0.98
$< 4, \mathbf{G}, A, sh >$	E^*_{avg}/L	0.09	0.21	0.278
$\lvert RSS_{th}\rvert = 3$	d^*/L	0.95	0.97	0.98
$< 4, \mathbf{R}, A, sh >$	E^*_{avg}/L	0.19	0.22	0.276
	d^*/L	0.425	0.376	0.355

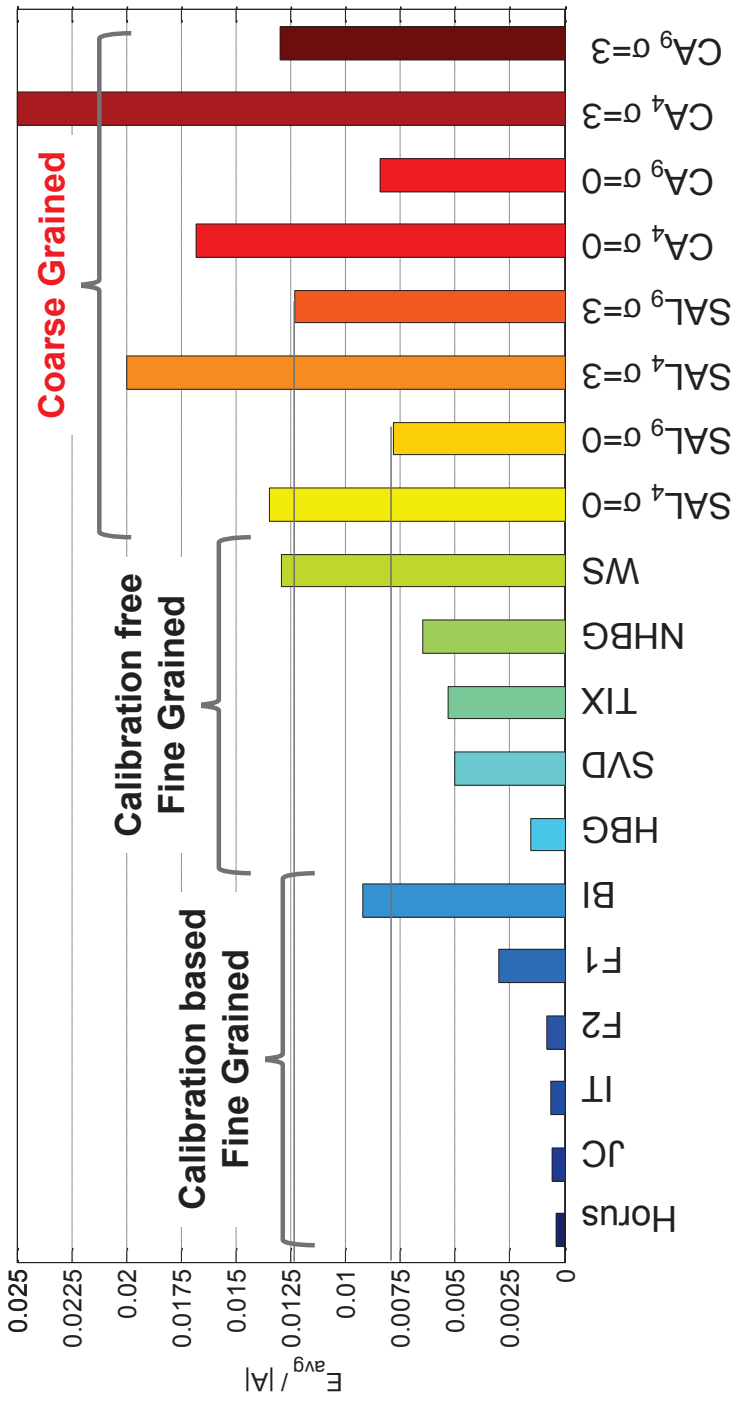

Figure 3.55: How does SAL compare to fine grained localization?

4.0 MULTI-FLOOR SUB-AREA LOCALIZATION

In Chapter 3, we proposed a localization approach that used the sub-areas created by overlapping coverages of MoSs placed on a virtual grid on a two dimensional area. We showed that this approach, called Sub Area Localization (SAL) is significantly better in terms of accuracy than using the centroids of the MoSs hearing a mobile node (MN). We investigated the relationship between localization accuracy and the number/monitoring range of MoSs. We employed analysis and simulations that allowed us to determine the "best" monitoring range of MoSs, and limits on the performance of this approach in 2-D environments. We showed that accuracy achieved at the best monitoring range follows an exponential improvement with increasing number of MoSs placed on a grid. This leads to a trade off as increasing number of MoSs also increases the cost as more MoSs have to be deployed. SAL is a proximity based localization algorithm that is suitable for single floor environments, as it is developed for two dimensional areas which can be considered as single floor areas. When a SAL system is deployed on a multi-floor building, for location estimation on each floor, it only utilizes MoSs that are located on *that* single floor.

Multi-floored environments provide us an opportunity to exploit the coverages of MoSs that are across floors. In this chapter, we are interested in the possibility and advantages of utilizing a proximity-based multi-floor localization approach. We are proposing *Multi-floor Sub-Area Localization (MSAL)* which can exploit MoSs across floors if they can sense transmissions across floors to achieve a satisfying accuracy with reduced deployment cost. To the best of our knowledge, this is the first work proposing to exploit MoSs in adjacent floors to estimate the location of a MN on the current floor. An obvious difference between the work presented in literature and this chapter is that we evaluate whether exploiting MoSs deployed in other floors to locate a node in the current floor can improve localization

accuracy without increasing cost. We use only the proximity information for localization. No systematic study was previously done on the effect of number of MoSs per floor, and on the effect of monitoring range and floor height while utilizing MoSs across floors.

1. We first evaluate whether exploiting MoSs deployed across floors is beneficial in terms of accuracy and efficiency. Our efficiency definition, as in the case of SAL, comprises good accuracy, high availability, and utilizing as few MoSs as possible per floor in a multi-floor building. We provide the results of a broad investigation of the effect of monitoring range on localization performance under various localization scenarios and channel conditions. We define a localization scenario as a localization algorithm using a given number of MoSs on each floor i, N_i, and placement ($\mathbf{P}$) of MoSs in a given localization area per floor ($A = L \times L$) in a building of F floors, and denote as $< (N_1, N_2, \ldots, N_F), \mathbf{P}, A >$.
2. We define the antenna type of MoSs, the relationship between floor height and monitoring ranges, the repetition/placement scheme employed as parameters needed to be considered through a systematic analysis of efficiency of MSAL. We investigate the difference in performance with isotropic antenna pattern and modified dipole antenna pattern. The ideal isotropic antenna pattern shows better accuracy as an outcome of its wider coverage in the vertical direction. Under ideal conditions, we show that by utilizing MoSs across floors, the localization error can be decreased by 40% without increasing the total number of MoSs when compared to single floor SAL.
3. We investigate the effect of log-normal shadowing, and show that under shadowing with 5 dB standard deviation, MSAL can reduce the cost of localization system by one third of the number needed if SAL was used, and with only a 15cm reduction in accuracy in a building with $10m \times 10m$ floor area, and floor height of 2.5m. Under ideal conditions, the accuracy gets worse only by 4cm. We have also investigated the performance of multi-floor localization when signal power is attenuated while passing through walls or traveling from other floors.
4. Optimization of MoS placement in multiple floors is very important to maximize the efficiency of localization system. In this chapter, rather than finding an optimized placement scheme, we provide the results of simulations of various placements, and show the performance differences between them. We have defined two and three floor repetition

schemes (2FRS and 3FRS) in which location estimates are calculated utilizing MoSs on two adjacent floors or three adjacent floors, respectively. With grid and random placement schemes, we have shown that 2FRS and 3FRS can outperform multi-floor localization scenarios utilizing centroids of MoSs or single-floor localization scenarios.

5. The number of MoSs deployed per floor is a very important parameter that has impact on the efficiency of localization. We have shown that using the same total number of MoSs on three floors, we can get a better efficiency when compared to using the same total number of MoSs on two floors.

6. In the first part of the analysis, we are focused on the advantages of utilizing MoSs across floors, therefore, we do not consider floor determination issues with MSAL. We assume that an external floor determination mechanism is available to estimate which floor an MN is located at.

7. In the second part, we propose that a majority-based floor determination algorithm using RSS readings from different floors can be incorporated into MSAL. We provide the floor determination performance under various channel conditions. We define the most important challenge related to floor determination as minimizing false negatives (the number of MNs that are located on a floor that they are not residing upon) with appropriate MoS placement. Our objective is to investigate the effect of floor determination on the performance of MSAL, rather than finding the best placement/number of MoSs that is optimizing the floor determination parameters and maximizing its performance.

Section 4.1 describes the multi-floor sub-area localization and MoS deployment schemes. Section 4.2 explains the multi-floor centroid algorithm (MCA) that we propose in order to make a fair comparison between MSAL and another multi-floor proximity based algorithm. We evaluate the performance of the MSAL in utilizing MoSs across floors in detail in Section 4.3 without considering floor determination effects. Then, in Sections 4.4 and 4.5, we present the majority based floor determination algorithm and its effects on the performance of MSAL with 2FRS localization scenario. Similar results with 3FRS are presented in Appendix A.3 for interested reader. Finally, Section 4.6 concludes the chapter with a set of tables summarizing our findings.

4.1 MULTI-FLOOR SUB-AREA LOCALIZATION (MSAL)

Localization error with SAL can be effectively decreased by optimizing the monitoring range for a given localization scenario, and can be further decreased by increasing the number of MoSs which would reflect as a higher cost localization system. Multi-floor Sub-Area Localization (MSAL) approach can exploit MoSs across floors if they can sense transmissions across floors to achieve a satisfying accuracy with smaller number of MoSs per floor. The target is reduced deployment cost without significantly reducing the accuracy. Multi-floor localization introduces a number of challenges and additional dimensions to the localization problem. Some of these are floor height, antenna coverage (monitoring range) across floors, floor repetition schemes, and the number/placement of MoSs, and channel effects.

A MN's location is denoted by a tuple $< x, y, f >$ representing 2 dimensional coordinates of a MN (x, y) on floor f. The floor f a MN is residing in is determined by an external floor determination algorithm that we assume available to the localization system. The focus of this chapter is to determine the performance of utilizing MoSs across floors rather than floor classification. We show how MoSs in multiple floors can be used with SAL to estimate the 2-D coordinates of a MN in a given floor. This multi-floor SAL algorithm is called MSAL in the rest of the chapter.

4.1.1 Adjacent Monitoring Range

On a single floor, the azimuthal coverage of MoSs is sufficient to determine the subareas (i.e., A_i in Figure 3.1) and thus the accuracy. In multi-floor localization, the propagation in the vertical direction gains importance. Antenna patterns and floor losses have to be considered to calculate monitoring ranges across floors. In what follows, we use an idealized antenna pattern (isotropic) and modified dipole antenna pattern; and ignore the impact of floor losses for now (although this could be captured by changing the floor height, h). Initially we assume an ideal channel without shadowing ($\sigma = 0$). Later, we allow σ of the normally distributed shadow fading component X to be non-zero.

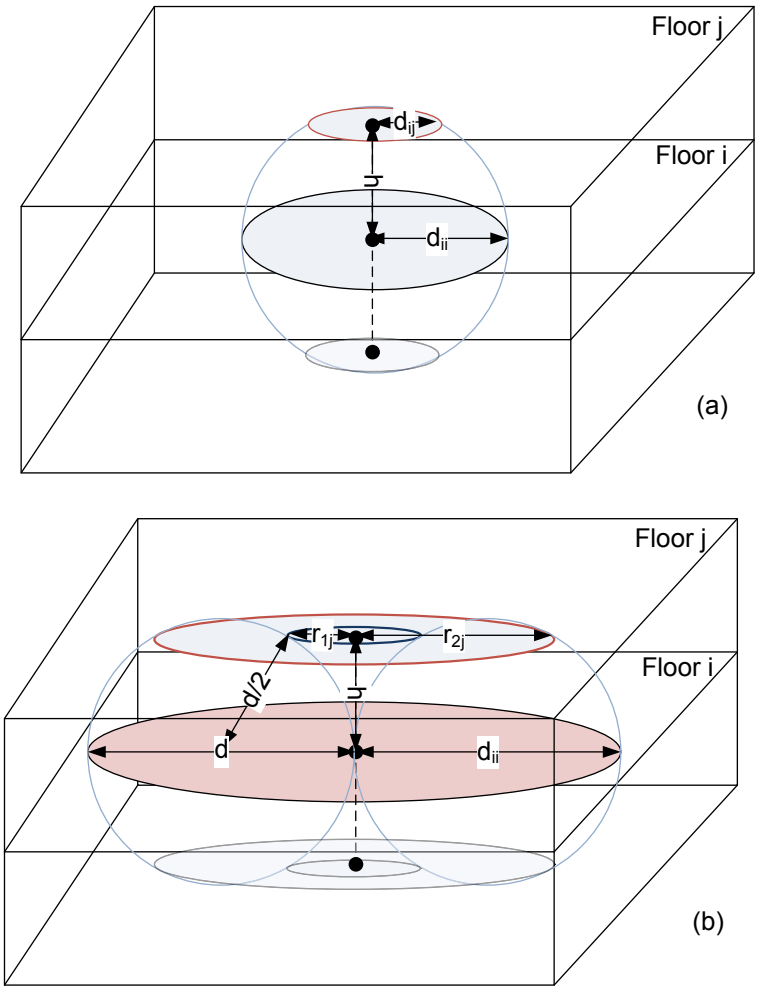

Figure 4.1: Range of an MoS in its own floor and in adjacent floors, (a) Isotropic antenna, (b) Dipole antenna

4.1.1.1 Using Isotropic Antenna Pattern

The monitoring range d of a MoS with an isotropic antenna will be the same in every direction. Let MoS_i be the set of MoSs deployed in floor i, h be the height of each floor, and d_{ij} be the effective monitoring range of MoSs in MoS_i on floor j. For localizing MNs on a given floor j, the effective monitoring ranges of MoS_i, d_{ij} for all $i = 1, \ldots, F$ can be exploited if a MoS can sense a MN in a different floor. A MoS on floor i has a smaller coverage or monitoring range in a different floor $j \neq i$. For an isotropic antenna, it is essentially a slice of a sphere that does not pass through the sphere's center. Given a MoS in MoS_i, the monitoring range in floor j is,

$$d_{ij} = \begin{cases} \sqrt{d_{ii}^2 - (|i-j|h)^2}, & d_{ii} > |i-j|h, \\ 0, & d_{ii} \leq |i-j|h, \end{cases} \tag{4.1}$$

where $i, j = 1, \ldots, F$ and F is the number of floors in a given building. The monitoring ranges of MoS_i in floor i and j are illustrated in Figure 4.1(a)[1] for adjacent floors when a MoS is placed in the center of a floor. We assume the monitoring range for all MoSs deployed on a single floor are the same. With MSAL and the use of MoSs on different floors, we end up with a situation which is similar to allowing different monitoring ranges for MoSs on a single floor. Figure 4.2 shows the effective monitoring range of a MoS on an adjacent floor with respect to monitoring range on current floor and floor height. Depending on floor height and monitoring range, MoSs may not cover any area on an adjacent floor. This happens when $d_{ii} < h$, and in this case only MoSs in the current floor can be utilized in localization. Therefore, there exists a range that is optimum for localization of MNs on both current floor and adjacent floor.

4.1.1.2 Using Modified Dipole Antenna Pattern

With a dipole antenna, monitoring ranges on the horizontal and vertical directions differ. This results in MoSs to have even smaller and ring shaped coverage on adjacent floors instead of circle shaped coverages as with isotropic antenna pattern. In Figure 4.1(b), the modified dipole antenna pattern is shown with its donut shaped coverage. We assume the cross section of the donut (when

[1]This analysis assumes that MNs are placed on the floor with zero height. Alternatively, this can be posed as a situation where the MoSs and MNs are at the same level on a given floor, and this level is defined as the "floor".

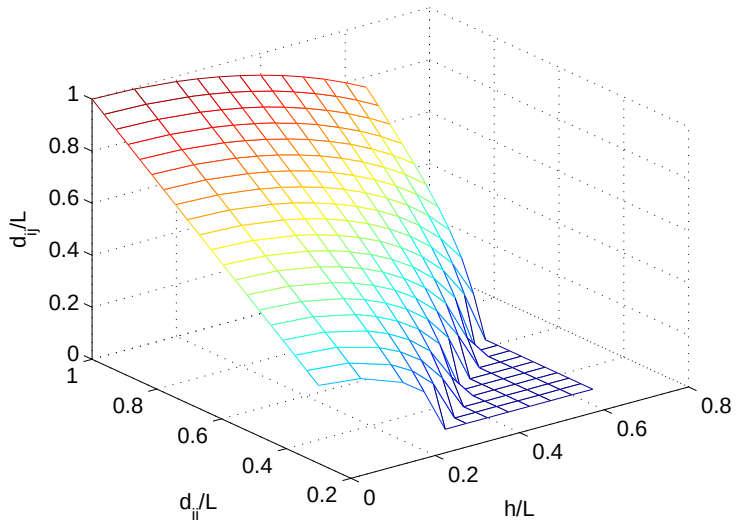

Figure 4.2: Range of an MoS in its own floor and in adjacent floors with Isotropic antenna

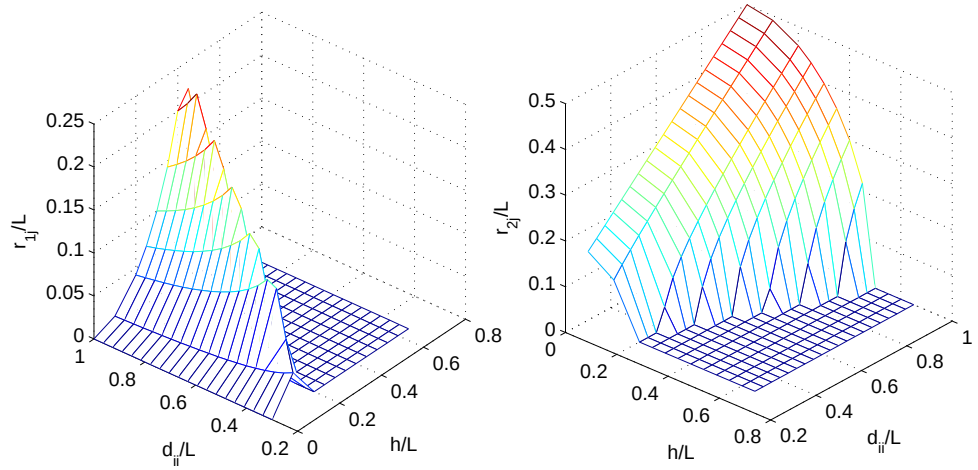

Figure 4.3: Range of an MoS in its own floor and in adjacent floors with Dipole antenna

sliced by a horizontal plane) is a circle for the sake of simplicity. We denote the effective monitoring range with modified dipole antenna as a tuple defining the radii of the smaller circle, r_{1j}, and larger circle of the ring, r_{2j}, on floor j, i.e., $d_{ij} = (r_{1j}, r_{2j})$. A MoS on floor i can have non-zero coverage on floor j if $d_{ii}/2 > |i-j|h$. When this condition is satisfied, $d_{ij} = (r_{1j}, r_{2j})$ are given as:

$$r_{1j} = \frac{d_{ii}}{2} - \sqrt{\left(\frac{d_{ii}}{2}\right)^2 - (|i-j|h)^2}, \tag{4.2}$$

$$r_{2j} = \frac{d_{ii}}{2} + \sqrt{\left(\frac{d_{ii}}{2}\right)^2 - (|i-j|h)^2}. \tag{4.3}$$

where $j = [max(i-1, 1), min(i+1, F)]$ in a building of F floors.

Figure 4.3 shows (r_{1j}, r_{2j}) with respect to the monitoring range on the current floor and the floor height. When compared to isotropic antennas, the adjacent floor monitoring range is much smaller with dipole antennas. We let the monitoring range on the current floor be the same as with isotropic antennas (in reality, this implies a reduction in the antenna gain), but the adjacent floor monitoring range is smaller in size and ring shaped. When the condition $d_{ii} < 2h$ is not satisfied, adjacent floor coverage is not possible. This complicates the process of finding the best monitoring range for both floors.

4.1.2 Utilizing Multi-Floor MoS Sets

Utilizing MoS sets deployed across floors is the idea behind reduction of deployment cost without harming accuracy with MSAL. In a multi-floor building, in order to achieve a required accuracy in each floor, MSAL requires a specified number of MoSs to be deployed on each floor. When only adjacent floors are considered, MSAL can use the coverages of MoSs from at most three different floors; two adjacent and one current floor. This scheme where all 3 available floors are used, and *repeated* through the whole building is called 3 floor repetition scheme (3FRS). If only one adjacent floor is utilized in addition to the current floor, it is called the 2 floor repetition scheme (2FRS). In order to demonstrate the advantage of MSAL over SAL, we compare MSAL with 2FRS and 3FRS scenarios to a localization scenario with SAL using the same total number of MoSs in each scenario. We assume that

the corresponding SAL localization scenario uses 9 MoSs in grid placement in 2D in area A, $< 9, \mathbf{G}, A >$ (evaluated in Chapter 3 with SAL). For MSAL, a deployment scheme is adopted so that the projection of MoSs onto one of the floors is similar to the deployment of 9 MoS on a grid in 2D. We call this placement scheme projected grid ($\mathbf{G}_p$).

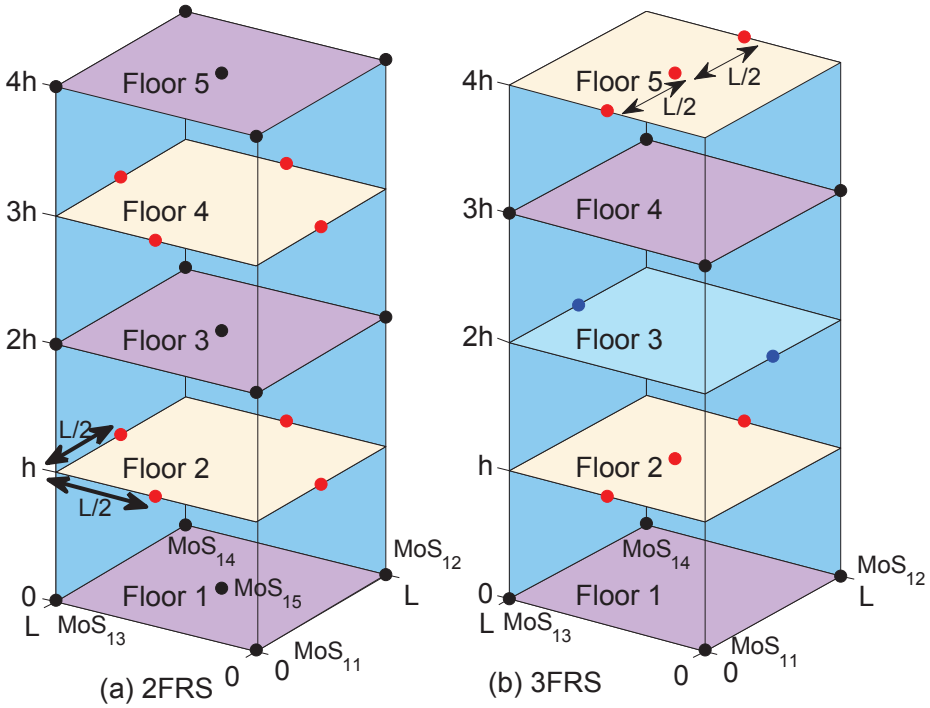

Figure 4.4: (a) 3FRS, $< (4,3,2), \mathbf{G}_p, A >$, (b) 2FRS, $< (5,4), \mathbf{G}_p, A >$

4.1.2.1 2FRS ($< (5,4), \mathbf{G}_p, A >$) The first MSAL scenario uses a 2 floor repetition scheme with 5 and 4 MoSs in the first and second floors with a "projected grid" placement scheme. The pattern of placement is shown in Figure 4.4(a). On one floor, MoSs are placed at the corners and also in the center of the floor, shown with black dots. In the adjacent floor, 4 other MoSs are placed at the centers of each edge of the floor boundary, shown with red dots. Differences exist in the creation of subareas between MSAL and SAL. Figures 4.5 (a) and (b) show the projections of the coverages of the 9 MoSs with isotropic antenna in each of the two floors with 4 and 5 MoSs with MSAL while Figures 4.5(c) and (d) do the same for only 4 or 5 MoSs all with the same range in a single floor with SAL. Here, we

used $d_{11} = d_{22} = 0.48L$, $h = 0.25L$ for MSAL, and $d = 0.48L$ for SAL. Visually, we can see that MSAL results in a larger number K of subareas with smaller sizes, thereby potentially improving the accuracy.

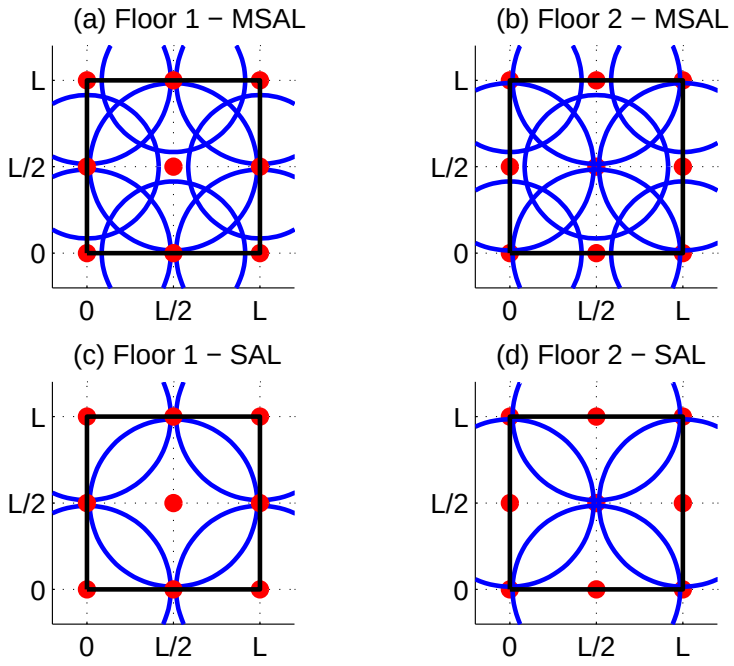

Figure 4.5: Subareas on floors 1 and 2 with MSAL $< (5,4), \mathbf{G}_p, A >$ and SAL $< 5, \mathbf{G}, A >$ and SAL $< 4, \mathbf{G}, A >$

In order to ensure 100% coverage with isotropic antennas on both floors (a MN will be detected by at least one MoSs on one of the two floors) according to unit disk model, following conditions must be satisfied:

$$\sqrt{d_{ii}^2 - \frac{L^2}{8}} + d_{ji} \geq \frac{\sqrt{2}L}{4},\ j \neq i,\ d_{ii} \geq \frac{\sqrt{2}L}{4} \tag{4.4}$$

$$\sqrt{d_{ji}^2 - \frac{L^2}{8}} + d_{ii} \geq \frac{\sqrt{2}L}{4},\ j \neq i,\ d_{ii} < \frac{\sqrt{2}L}{4} \tag{4.5}$$

We don't formally prove the above equation, but qualitatively show why this is true in Figure 4.6. In this figure, we show the monitoring ranges d_{ii} of MoSs on the same floor (larger circles) and those (d_{ji}) of the MoSs on the adjacent floor (smaller circles) for barely

covering Floor i. Case 1 in Figure 4.6(a) represents the 1st condition given in Eq. 4.4, and case 2 in Figure 4.6(b) represents the 2nd condition given in Eq. 4.5.

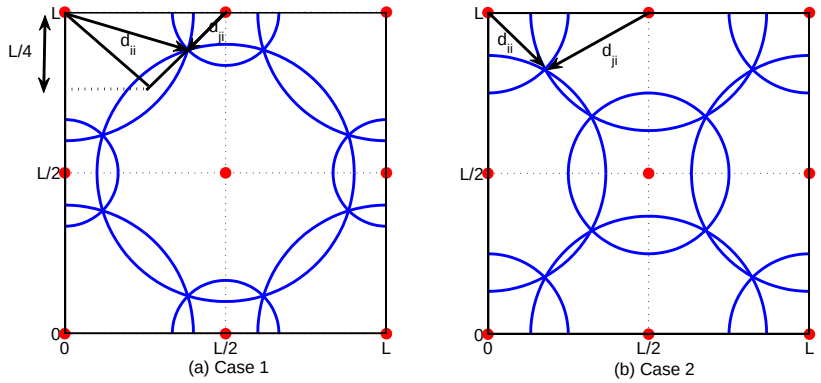

Figure 4.6: Ranges of MoSs in two floors for barely covering a floor.

4.1.2.2 3FRS ($< (4, 3, 2), \mathbf{G}_p, A >$) The second MSAL scenario uses a 3 floor repetition scheme with a total of 9 MoSs in projected grid placement. We place 4, 3, and 2 MoSs on adjacent floors in a building as shown in Figure 4.4(b). Instead of using 4.5 MoSs per floor with 2FRS, 3FRS uses 3 MoSs per floor to reduce the deployment costs per floor. The coverages of MoSs or their projections onto a single floor are shown in Figure 4.7. It is clearly seen that whole floor coverage is not available with SAL whereas with multi-floor MoS usage, full floor coverage is provided as well as higher number of subareas with small sizes without increasing the total number of MoSs deployed.

4.1.3 Availability

In Section 3.1.2, we defined availability as the fraction of area that a location estimate is available for a MN in localization area A. Let C_i' be the area on floor i that can be sensed by at least one MoS across floors. When $\sigma = 0$, MNs are always located in valid sub-areas whose boundaries are determined from the unit disk model. When $\sigma > 0$, an area $C_{invalid,i}$ of the localizable area C_i' may belong to invalid subareas. These arise because radio propagation vagaries may create intersections of monitoring areas of MoSs, that theoretically do not

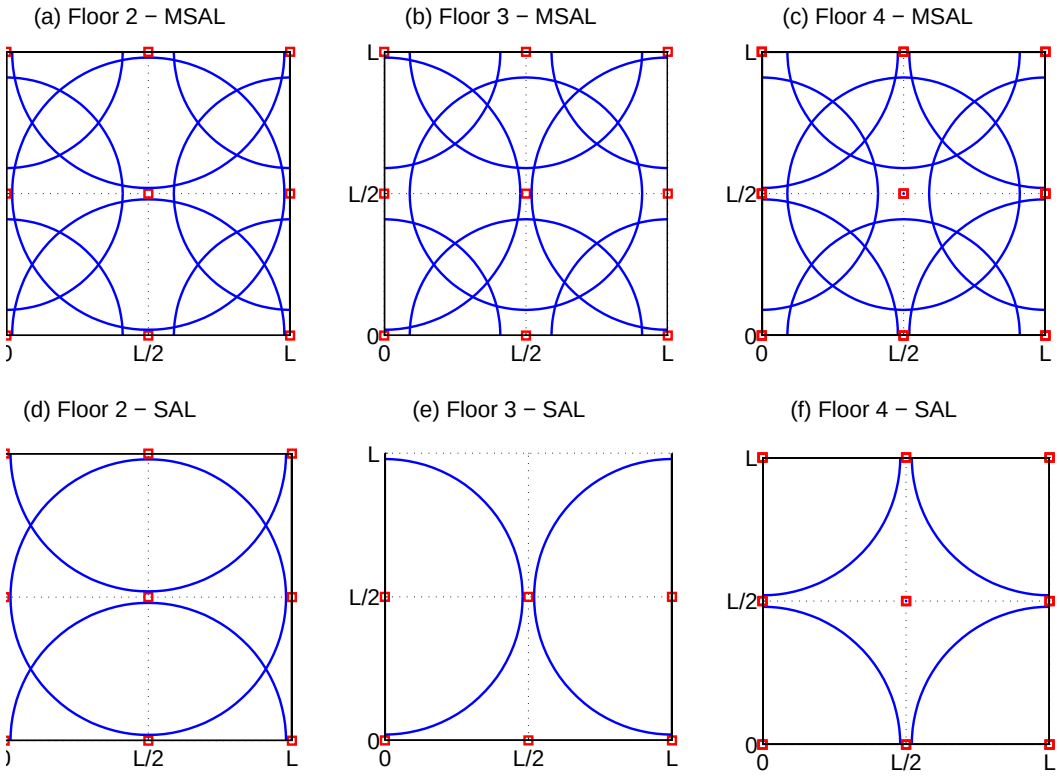

Figure 4.7: Subareas on floors 2,3 and 4 with MSAL $< (4,3,2), \mathbf{G}_p, A >$ and SAL $< 3, \mathbf{G}, A >$, SAL $< 2, \mathbf{G}, A >$, and SAL $< 4, \mathbf{G}, A >$

have intersecting ranges. MSAL cannot be directly used for localizing nodes in these invalid subareas. Nodes in such invalid subareas can be localized for instance, by using the multi-floor centroid algorithm that is explained in Section 4.2. Let $C_i = C'_i - C_{invalid,i}$. Then, the fraction of the area in which a node can be localized by MSAL on a floor i is

$$R_i = \frac{C_i}{|A|} \tag{4.6}$$

Then, on the average, a location estimate with MSAL is available for the fraction of area $R_{avg} = \frac{1}{F}\sum_{i=1}^{F} R_i$ in a building of F floors.

4.1.4 Efficiency

In order to determine the parameters so that MSAL can (i) be able to localize as large a fraction of the area as possible (ii) provide the smallest possible localization errors and (iii) use the smallest number of MoSs, we define a per-floor efficiency metric denoted by η_f. This metric evaluates the performance of MSAL on floor f for a given σ, h, d_{ff}, $f = 1 \ldots F$. η_f is directly proportional to the availability fraction, R_f, and inversely proportional to E_f, the mean error that can be achieved for such nodes in floor f with area $A = L \times L$, multiplied by $N_f = |MoS_f|$ – the number of MoSs deployed on floor f. Thus, per-floor efficiency metric is given as,

$$\eta_f(i,j) = \frac{R_f}{N_f^i \times (E_f/L)^j}, \tag{4.7}$$

and average efficiency over all floors is $\eta_{avg} = \frac{1}{F}\sum_{f=1}^{F} \eta_f$. In the following, we assume $i = j = 1$.

4.2 MULTI-FLOOR CENTROID ALGORITHM (MCA)

In Chapter 3, we compared the performance of proximity based SAL algorithm to the performance of CA which is another proximity based algorithm. In this chapter, in order to make a fair comparison between proximity based MSAL and another multi-floor proximity based algorithm, we propose multi-floor centroid algorithm (MCA). MCA works similar to centroid algorithm in that it also estimates the location as the centroid of the locations of MoSs hearing a MN. However, as a difference from CA, MCA utilizes MoSs across floors. In location estimation, only 2D coordinates of MoSs on multiple floors are taken into consideration. Effective monitoring ranges are calculated as given in Section 4.1.1.

4.3 ANALYSIS AND EVALUATION OF MSAL

In this section, we evaluate the two deployment schemes (2FRS and 3FRS) using isotropic or dipole antenna. We append the words *iso* and *dpl* to the algorithm names to indicate the isotropic and dipole antenna types are used, respectively. We determine the best monitoring ranges and best cost/accuracy metrics in each case and discuss the findings. While a uniform distribution of MNs can be assumed, there is no easy closed form mathematical expression for the localization error and we use simulations to determine the localization errors. We place a large number $M = 10^4 L$ MNs in the given area, and compute the error between the true location and the centroid of subareas (the estimated location) created by MoSs for each of the large number of MNs. The simulations are run for various monitoring ranges, heights of floors, deployment schemes, etc. as described below. We assume path loss coefficient $\alpha = 3$ and $P_t = 10dBm$.

4.3.1 Performance under Ideal Channel Conditions

We start our analysis by assuming ideal channel conditions. We compare MSAL performance in terms of various metrics to the performance of multi-floor centroid algorithm (MCA).

Again, we assume that the effective monitoring ranges of MoSs are known to the localization system which assumes a unit disk monitoring range model.

4.3.1.1 2FRS ($< (5,4), \mathbf{G}_p, A >$) We first investigate the performance of MSAL under localization scenario $< (5,4), \mathbf{G}_p, A >$ explained in Section 4.1.2.1, and shown in Figure 4.4(a) in a building of 5 floors. The results in this section are given for two floors only, but they can be generalized for $F > 2$ assuming that the range of MoSs on the odd numbered floors are the same and those on even numbered floors are the same. This ensures the subareas do not change from what we have for $F = 2$. Therefore, the minimum achievable errors and maximum efficiency will stay the same when $F > 2$.

Figure 4.8 shows the average expected errors separately in floors 1 & 2, E_1 and E_2, and average error for both floors, $E_{avg} = \frac{1}{2}\sum_{i=1}^{2} E_i$, for different values of monitoring ranges, $d_{ii},\ i = 1, 2$. Floor height is assumed $h = 0.25L$. The first column of figures shows MSAL-iso, and the second column shows MSAL-dpl localization error. Only the set MoS_i is used for localization in floor i, when $d_{jj} \leq h$ in isotropic case, and $d_{jj} \leq 2h$ in dipole case. Under these conditions, MSAL is equivalent to SAL. This is the reason why we are observing constant error on floor 1 when $d_{22} \leq 0.5L$ when $h = 0.25L$, and similarly on floor 2 when $d_{11} \leq 0.5L$ with MSAL-dpl.

We define accuracy as the minimum average localization error that can be achieved when the whole area A is covered by MoSs, i.e., $R = 1$ on every floor under ideal conditions. It is denoted as E^*_{avg},

$$E^*_{avg} = min_{\substack{d_{ii}, i=1\ldots F \\ s.t.\ R_i=1}} \frac{1}{F}\sum_{i=1}^{F} E_i \tag{4.8}$$

$$d^*_{ii} = argmin_{d_{ii}, i=1\ldots F\ s.t.\ R_{MSAL}=1} \frac{1}{F}\sum_{i=1}^{F} E_{avg},\ i = 1\ldots F. \tag{4.9}$$

and the monitoring ranges that provide this accuracy are denoted as d^*_{ii}. Table 4.1 summarizes the minimum localization error conditions with SAL, MSAL-iso and MSAL-dpl. According to this table, **by utilizing MoSs across floors, the localization error can be decreased by 40% without increasing the total number of MoSs**. This concludes that under ideal conditions, when best monitoring ranges are determined, isotropic

and dipole antenna coverages lead to similar localization error performances using MSAL. With MSAL-dpl, the best monitoring ranges are longer than needed for MSAL-iso. Another observation is that the accuracy with MSAL is similar to the accuracy ($0.078L$) achieved when SAL is used with $N = 9$ MoSs with scenario $< 9, \mathbf{G}, A >$ as shown in Chapter 3.

Table 4.1: Minimum localization error in 2FRS, $\sigma = 0$

	E_1/L	E_2/L	E^*_{avg}/L	$(d^*_{11}, d^*_{22})/L$
SAL (5, 4)	0.135	0.138	0.136	(0.77, 0.74)
MSAL-iso	0.0807	0.0814	0.081	(0.85, 0.70)
MSAL-dpl	0.0823	0.0801	0.0812	(0.88, 0.74)
SAL (9, 9)	0.078	0.078	0.078	(0.45, 0.45)

Figure 4.9 shows the various subareas and estimated locations for each subarea on each floor when the accuracy is $0.081L$ (81 cm accuracy in a 100 square meter area) with MSAL-iso. This accuracy is achieved when the ratio of the monitoring ranges on adjacent floors is $d_{11}/d_{22} = 1.2$. Intuitively, one would expect this ratio to be closer to 1, but the asymmetry in deployment and floor height make it different from 1. The red line in Figure 4.8 shows where $d_{11} = d_{22}$ and E^*_{avg} on this line is $0.082L$ which is close to E^*_{avg}, at $d_{11} = d_{22} = 0.778L$.

We emphasize that the minimum localization error values mentioned are determined only when the availability of location estimation is 100%. In order to better understand this statement, we demonstrate the availability with MSAL-iso and MSAL-dpl when $h = 0.25L$ in Figure 4.10. As expected a greater range of monitoring ranges can provide 100% availability with isotropic antennas due to wider adjacent monitoring range. Because of ring shaped adjacent coverages, with dipole antennas, each d_{ii} has to ensure full coverage on floor i separately. Then, with scenario $< (5, 4), \mathbf{G}_p, A >$, if $d_{ii} \geq 0.5L$, $i = 1, 2$, availability is 100%. Note that under ideal channel conditions, $C_{invalid,i} = 0$. In what follows, we only consider nodes with available location estimations, and provide the localization error only for these nodes.

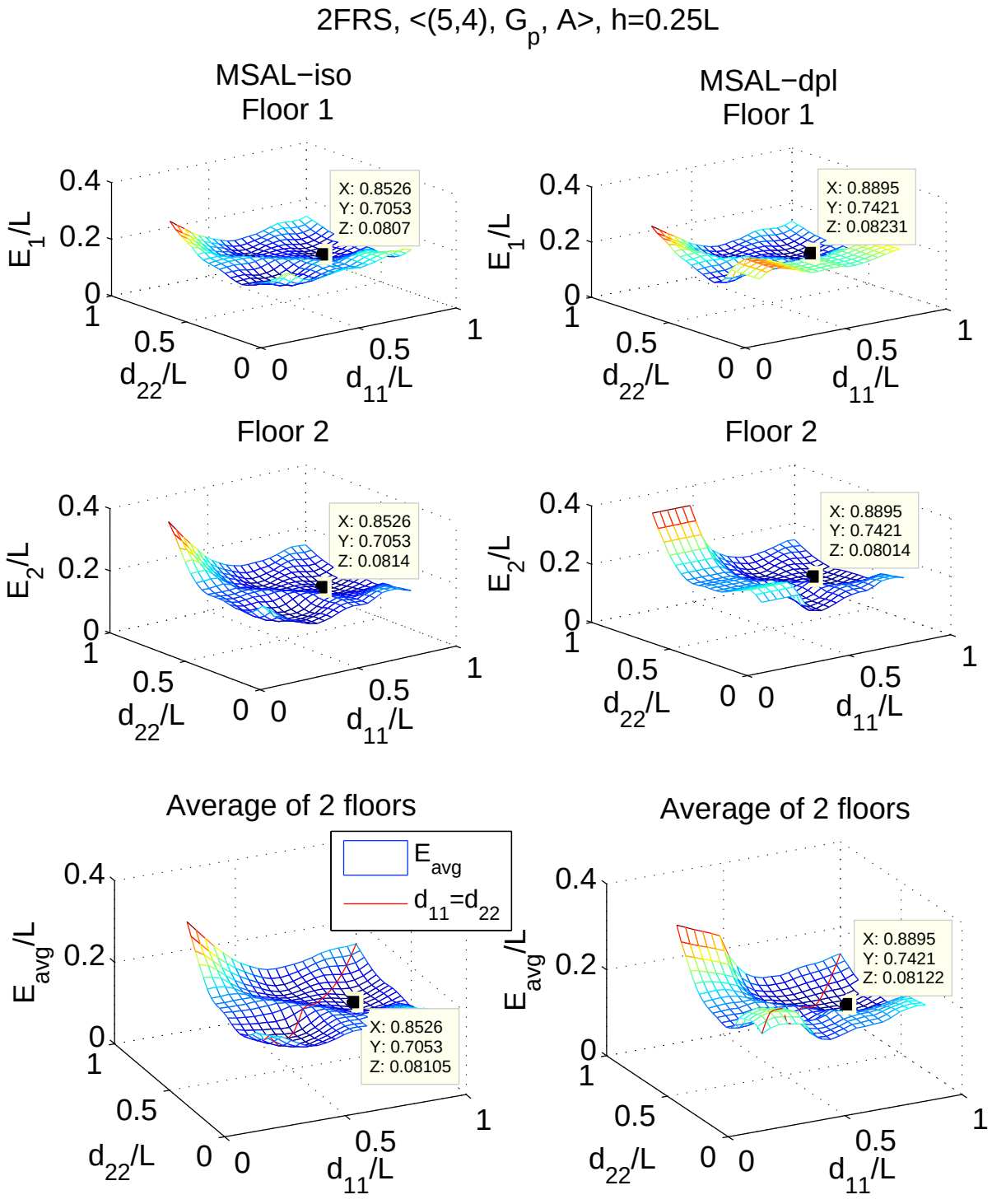

Figure 4.8: Localization errors on floors 1 and 2 with MSAL-iso and MSAL-dpl, $\sigma = 0$, $h = 0.25L$

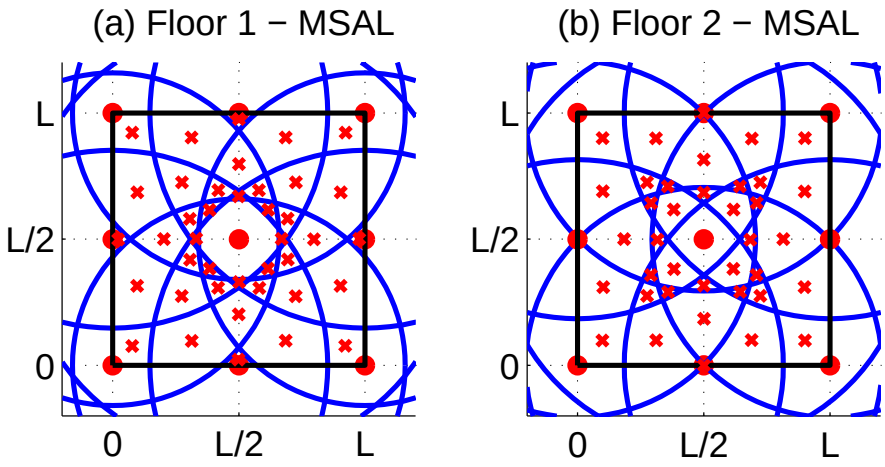

Figure 4.9: Sub-areas in floors 1 and 2 for $E^*_{avg} = 0.081L$, $\sigma = 0$

Figures 4.11 and 4.12 show the *minimum average error* and the corresponding *best* monitoring ranges on two floors with respect to h for MSAL and MCA. We observe that when all MoSs are deployed in one floor ($h = 0$) as in SAL with $< 9, \mathbf{G}, A >$, the minimum mean error is approximately $0.078L$ with both dipole and isotropic antenna. Note that in this case we allow two different monitoring ranges for 2 sets of MoSs and we observe the best monitoring ranges for two sets are the same, i.e., $d^*_{11} = d^*_{22}$ due to current floor coverages for isotropic and dipole antennas having the same size and shape. When $h = 0$, MSAL is equivalent to SAL, and MCA is equivalent to CA. When $h > 0$, MSAL and MCA can be employed. As h is increasing, d^*_{ii} must be increased in order to utilize adjacent floor monitoring ranges. The accuracy with MSAL-iso gets worse with increasing h as the effective monitoring ranges on adjacent floors get smaller. With dipole antennas, when $h < 0.25L$, the accuracy is similar to the isotropic case. When h continues to increase we observe a slightly better accuracy due to a change in best monitoring range on floor 1. This is due to the creation of the same number of subareas each now represented by different set of MoSs. From $h = 0.25L$ to $h = 0.35L$, the number of subareas stays the same; however, the sizes

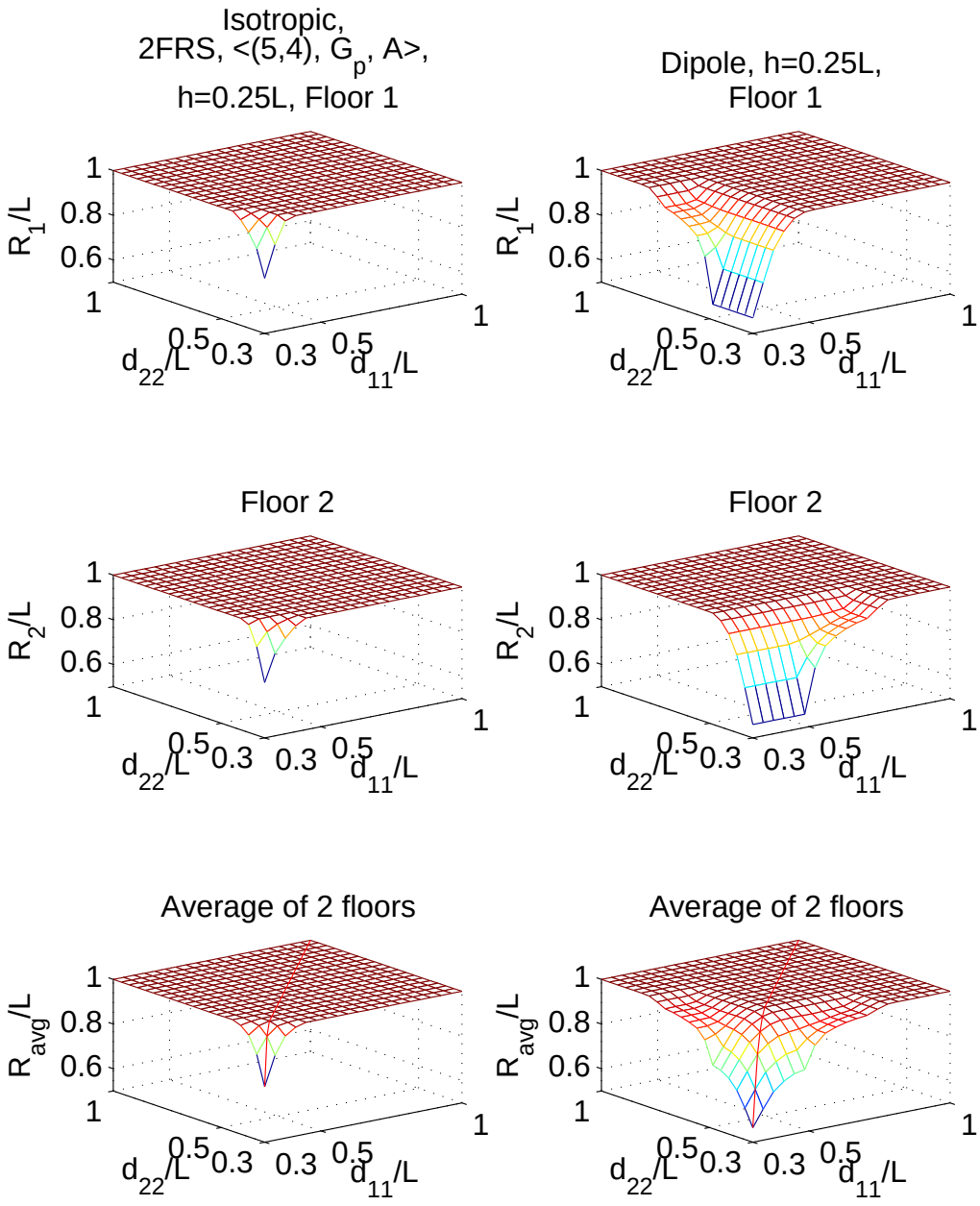

Figure 4.10: Effect of monitoring ranges and floor height on availability

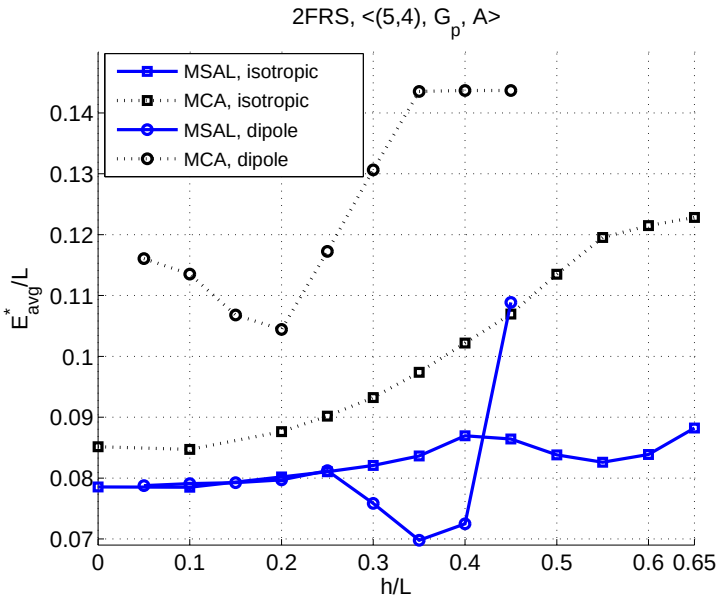

Figure 4.11: Effect of floor height on E^*_{avg} with 2FRS

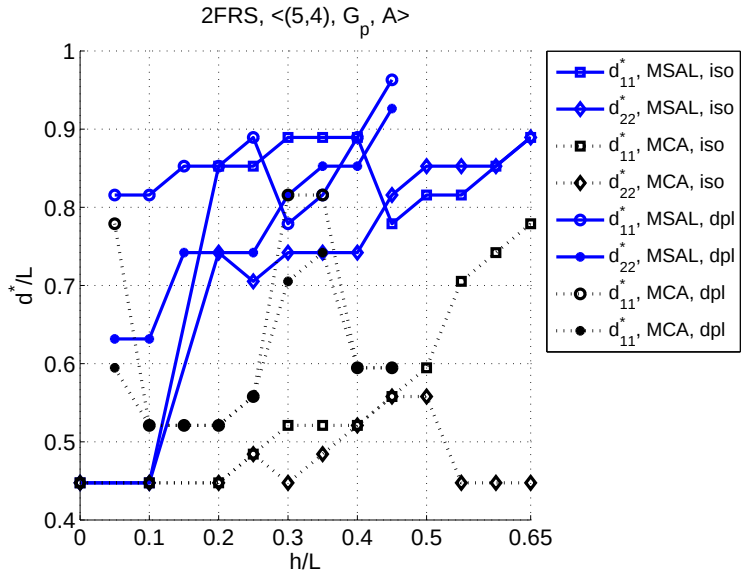

Figure 4.12: Best monitoring ranges $d^*_{ii}/L,\ i = 1, 2$ with 2FRS

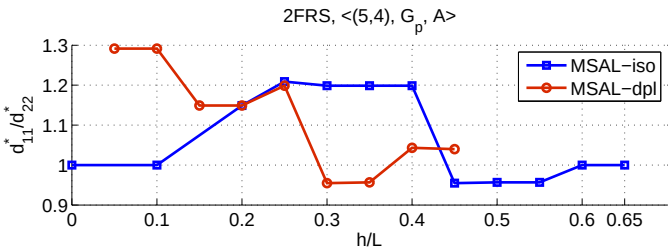

Figure 4.13: $\frac{d^*_{11}}{d^*_{22}}$ with 2FRS

are optimized and balanced with the best monitoring ranges that will allow better accuracy. When $h > 0.35L$, for 100% coverage on each floor, the monitoring ranges are increased. This is reflected in the creation of multiple subareas in different locations (symmetric) in A and represented however by the same set of MoSs. Since the localization system is not aware of this, the estimate is given as the center of A, and this effect worsens the accuracy when $h = 0.45L$.

We also observe that MCA has a worse accuracy than MSAL. It is important to notice that when $h > 0.35L$, MCA-dpl stops using multi-floor MoSs, and instead CA uses only the current floor MoSs for location estimates. This is due to the need to increase the monitoring ranges by a factor of 2 of h. If h is large, this will cause a very large monitoring range on the current floor, and not enough coverage on the adjacent floor. Then, CA will not be able to estimate locations as accurately as possible with smaller monitoring ranges.

Figure 4.13 shows the ratio of current floor range to adjacent MoS range. We observe a ratio close to 1. Also, as h increases the ratio gets closer to 1.

4.3.1.2 3FRS ($< (4,3,2), \mathbf{G}_p, A >$) We also investigate the performance of MSAL under a 3 floor repetition localization scenario $< (4,3,2), \mathbf{G}_p, A >$ explained in Section 4.1.2.1. In order to show the effect of this deployment scheme, we assume 4, 3, 2, 4, 3 MoSs are placed on the 1st, 2nd, 3rd, 4th and 5th floors in a 5-floor building as shown in Figure 4.4(b). In simulations, to avoid the complexity that will be introduced by an additional variable, we allow two different monitoring ranges to be set among 3 floors. This will also keep the

calibration efforts at a lower level as two different monitoring ranges for MoSs has to be calibrated instead of three. We assume floors with 4 and 3 MoSs are allowed to have the same monitoring range for all MoSs deployed on those floors, and floors with 2 MoSs are allowed to have a different monitoring range than the other floors have. Therefore, floors 1,2,4, and 5 have the same monitoring range d_{11}, and floor 3 has monitoring range d_{22}. The results in this section are given for three floors only – namely floor 2, 3 and 4 – in order to present the 3FRS performance.

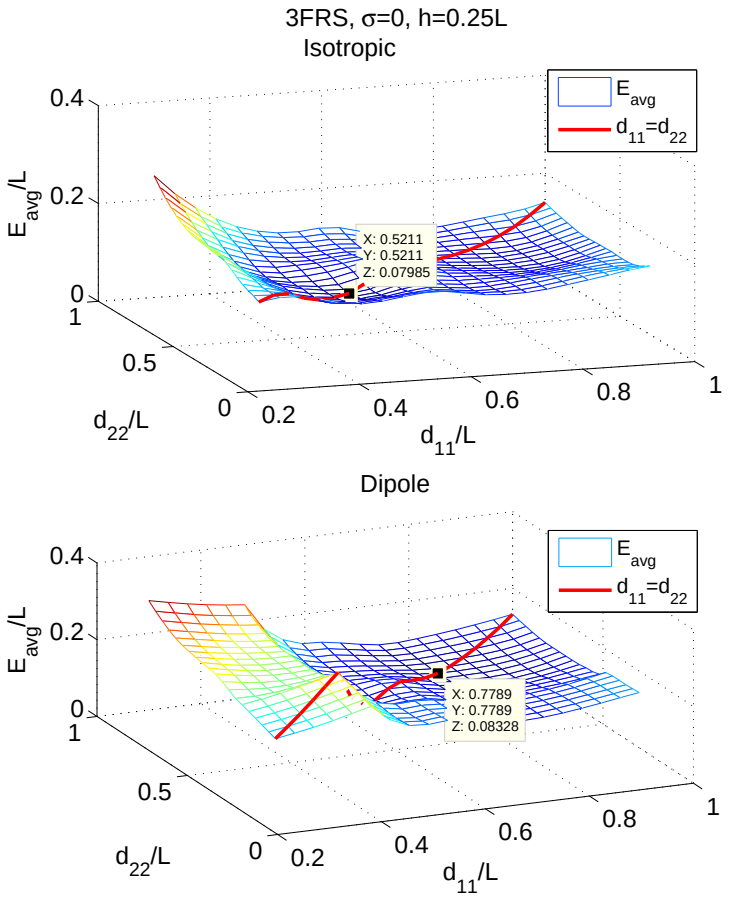

Figure 4.14: E_{avg} in 3FRS with isotropic and dipole antenna, $\sigma = 0$, $h = 0.25L$

Figure 4.14 shows the average expected error in 3 floors for different values of monitoring ranges, d_{ii}, $i = 1, 2$ when floor height is given as $h = 0.25L$ with MSAL-iso and MSAL-dpl. Table 4.2 shows E^*_{avg}, and corresponding errors on each floor using SAL and MSAL-

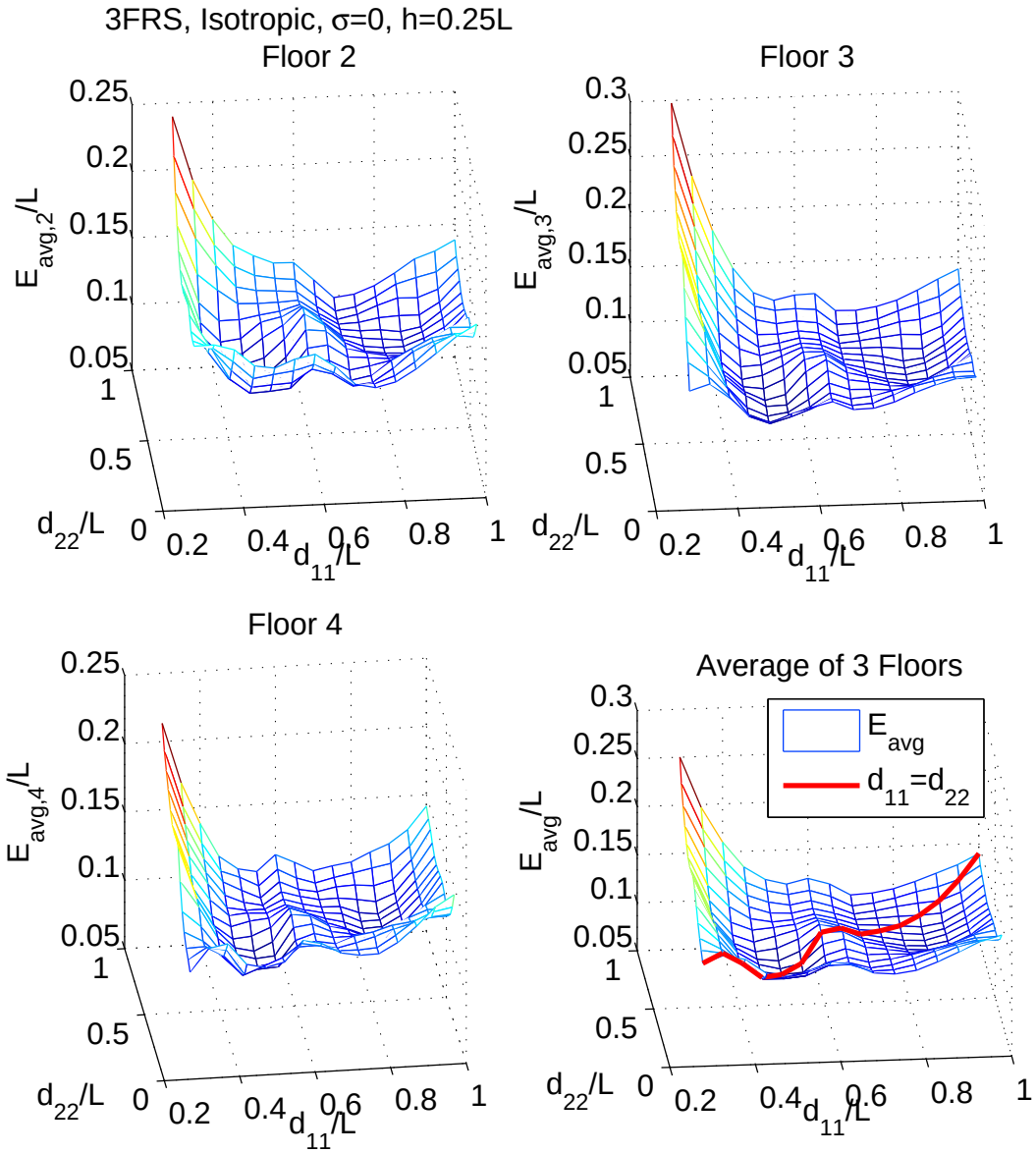

Figure 4.15: E_{avg} for 3FRS with isotropic antenna, $\sigma = 0$, $h = 0.25L$

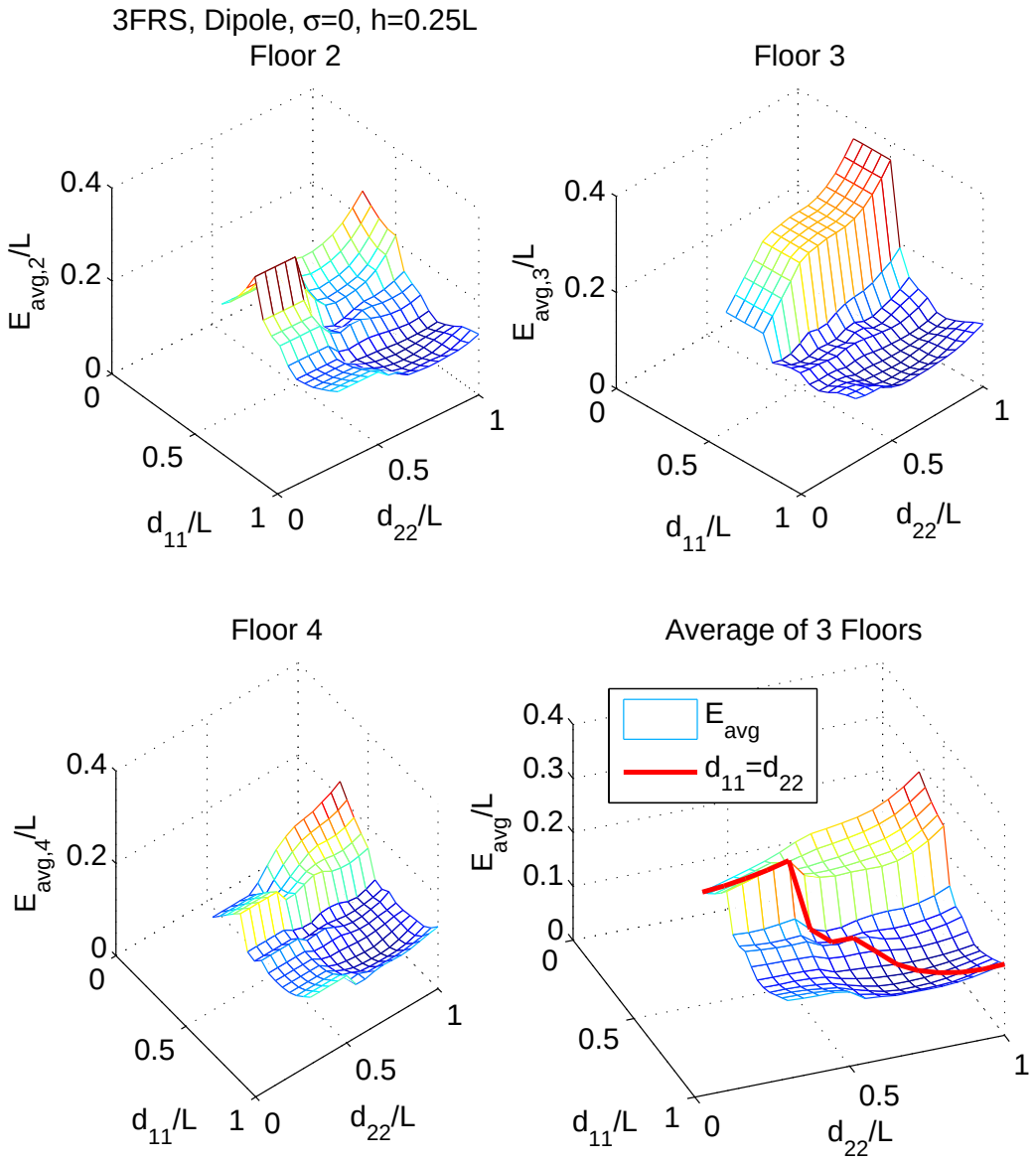

Figure 4.16: E_{avg} for 3FRS with dipole antenna, $\sigma = 0$, $h = 0.25L$

Table 4.2: Minimum localization error conditions in 3FRS

	E_1/L	E_2/L	E_3/L	E^*_{avg}/L	$(d_{11}, d_{22}, d_{33})/L$
SAL(4,3,2)	0.135	0.27	0.19	0.198	(0.74,0.56,0.71)
MSAL-iso	0.082	0.079	0.077	0.079	(0.521,0.521,0.521)
MSAL-dpl	0.086	0.085	0.077	0.083	(0.778,0.778,0.778)

iso and MSAL-dpl. Under ideal conditions, best monitoring ranges can lead to similar localization errors with isotropic and dipole antennas. The variation in accuracy between floors is negligible as shown in Figures 4.15, 4.16 and Table 4.2.

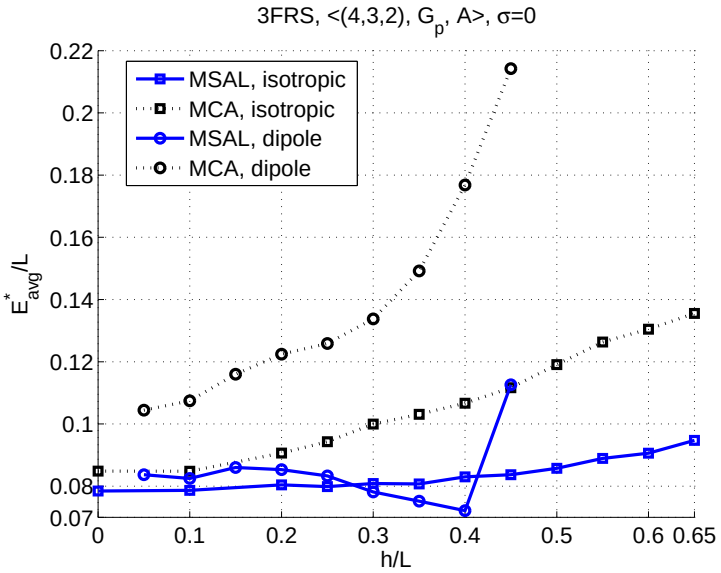

Figure 4.17: Effect of floor height on E^*_{avg} with 3FRS

Figure 4.17 shows the minimum achievable localization error with respect to floor height with MSAL-iso/dpl and MCA-iso/dpl. The behavior seen in MSAL-iso and MSAL-dpl are similar to the behavior seen in 2FRS scenario. MSAL-iso has an increasing trend, while MSAL-dpl first decreases, then starts increasing. MSAL outperforms MCA significantly

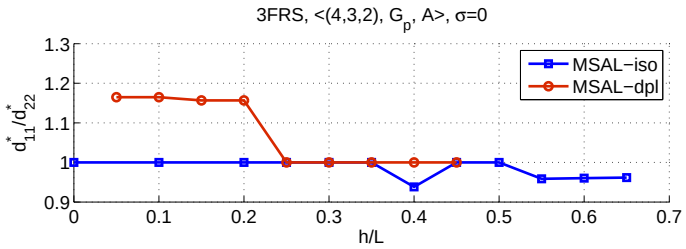

Figure 4.18: Best monitoring ranges $d^*_{ii}/L,\ i = 1, 2$ with 3FRS

especially when h is large, and dipole antennas are used. Figure 4.18 shows the ratio of the best monitoring ranges on two floors is close to 1.

4.3.2 Impact of Log-Normal Shadowing

Varying channel conditions has a significant impact on localization performance. The vagaries in channel may result in creation of subareas that are not known by MSAL localization system which is assuming unit disk model. These subareas are counted in $C_{i,invalid}$ during discussion of availability in the previous section. When $C_{i,invalid} > 0$, we propose to use MCA for MNs that are detected in these *invalid* areas, and MSAL for MNs detected in *valid* subareas. We call this algorithm MSAL+MCA. We expect that this algorithm will have higher availability and worse localization error than MSAL due to localization errors of MNs detected in invalid subareas just as in the case of SAL. These MNs' location estimations are not available when only MSAL is employed.

In this section, we investigate the effect of standard deviation in shadowing on location accuracy, availability and efficiency with MSAL, MCA and MSAL+MCA with 2FRS and 3FRS scenarios. We assume isotropic or dipole antennas are used.

Figures 4.19 and 4.20 show that when $h = 0.25L$, the localization accuracy gets worse under increasing shadowing effect with both 2FRS and 3FRS scenario. Other observations are as follows: (1) 3FRS has slightly worse accuracy (6.6% on the average) than 2FRS due to more MoSs with smaller adjacent ranges. (2) MSAL performs best under all shadowing levels; however, we remind that accuracy is given for MNs for which an estimate is available.

Table 4.3: MSAL acronyms and variables

Variable	Explanation
MoS_i	Monitoring station deployed on floor i
d_{ii}	Monitoring range on floor i of a MoS_i
d_{ij}	Effective monitoring range on floor j of a MoS_i
h	Floor height
2FRS	2 Floor repetition scheme
3FRS	3 Floor repetition scheme
E^*_{avg}	Average accuracy on multiple floors
η^*_{avg}	Maximum achievable average efficiency on multiple floors
R_{avg}	Average of availability on multiple floors
$f*$	Estimated floor according to MBFD algorithm
$A_{fn,i}$	The fraction of area on floor i in which MNs cannot be located on the correct floor.

Figures 4.21 and 4.22 show the availability under best accuracy conditions under different shadowing conditions. Table 4.4 summarizes the observations from these figures, and shows that availability with MSAL is significantly affected by shadowing when σ is large. Under severe channel conditions, MSAL+MCA can provide close to maximum availability. Figures 4.23 and 4.24 show the monitoring ranges with MSAL+MCA is decreasing as σ increases. The reason of higher availability at $\sigma = 10$ with MSAL-dpl is that at this shadowing level, SAL is used in order to minimize the localization error. When $h = 0.25L$, $d_{ii} > 0.5L$ for multi-floor localization has to be used.

Figures 4.25 and 4.26 shows the behavior of MSAL-iso and MSAL-dpl accuracy with

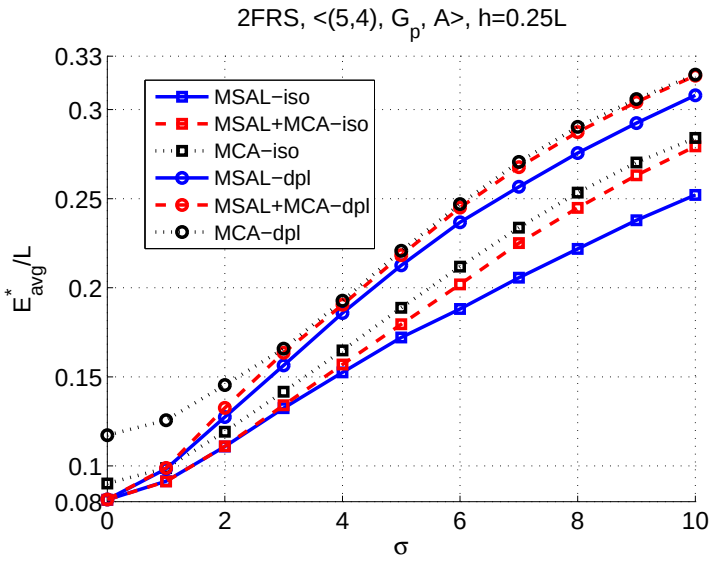

Figure 4.19: Effect of shadowing on E^*_{avg} with 2FRS

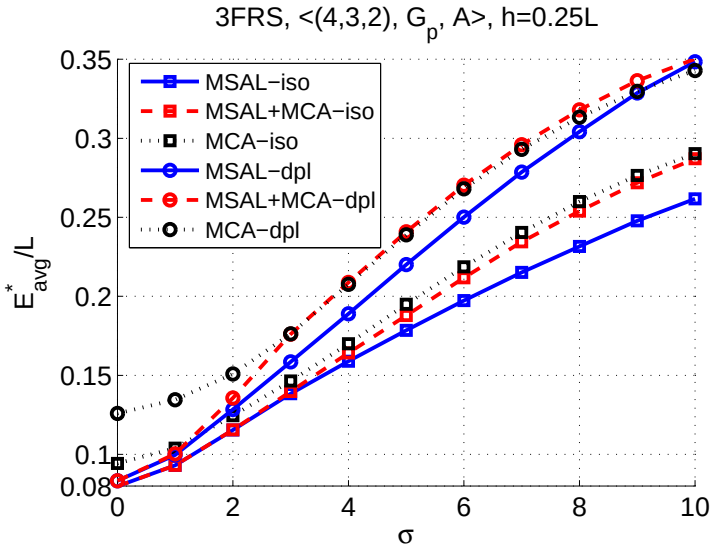

Figure 4.20: Effect of shadowing on E^*_{avg} with 3FRS

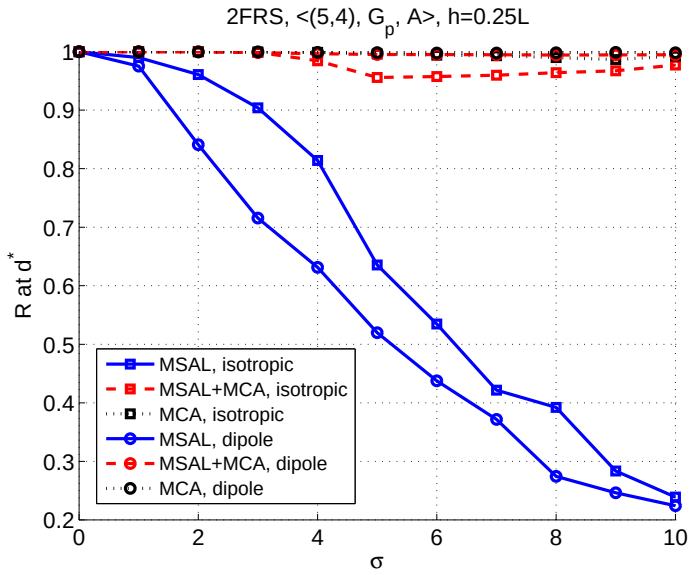

Figure 4.21: Availability at d^*_{ii} 2FRS with isotropic and dipole antenna, $h = 0.25L$

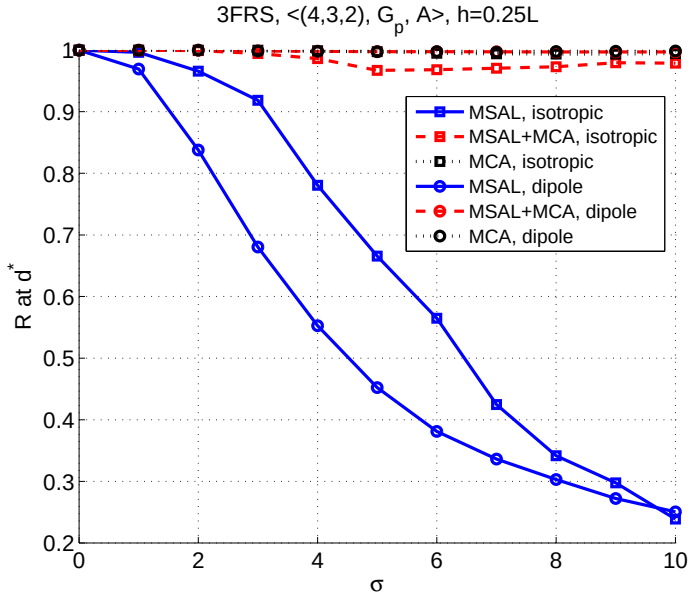

Figure 4.22: Availability at d^*_{ii}, 3FRS with isotropic and dipole antenna, $h = 0.25L$

Table 4.4: Availability at minimum error conditions ($h = 0.25L$)

	2,3FRS $\sigma = 0$	2FRS $\sigma = 5$	2FRS $\sigma = 10$	3FRS $\sigma = 5$	3FRS $\sigma = 10$
MSAL-iso	1	0.63	0.24	0.66	0.23
MSAL-dpl	1	0.40	0.58	0.38	0.45
MSAL+MCA-iso	1	0.96	0.98	0.96	0.97
MSAL+MCA-dpl	1	0.99	0.98	0.99	0.99
MCA-iso	1	0.99	0.99	0.99	0.99
MCA-dpl	1	0.99	0.99	0.99	0.99

2FRS and 3FRS under different floor height and shadowing conditions. The observation is in line with expectations as smaller floor height and smaller σ provide better accuracy in all cases, and as σ increases the effect of large floor height is more significant.

Figure 4.27 gives an idea on how MSAL accuracy performance compares to other localization algorithms under shadowing with $\sigma = 5$ for different floor heights. MSAL-iso outperforms other algorithms when location estimate is available. When compared to accuracy under ideal conditions given in Figure 4.11, we still observe increasing trend in localization error with isotropic antennas. We also observe increasing trend with dipole antennas as opposed to the behavior under ideal conditions. With isotropic antennas it is best to use MSAL+MCA in order to get as many estimates as possible (Table 4.4), and still less error than MCA. MCA-dpl outperforms MSAL+MCA-dpl if $h < 0.2L$; otherwise, MSAL+MCA-dpl is again the best choice. The observation of efficiency at minimum localization error conditions are in line with this conclusion.

In order to find out the overall efficiency performance of MSAL compared to other two algorithms, we look at Figure 4.28 showing the efficiency that is provided by the best mon-

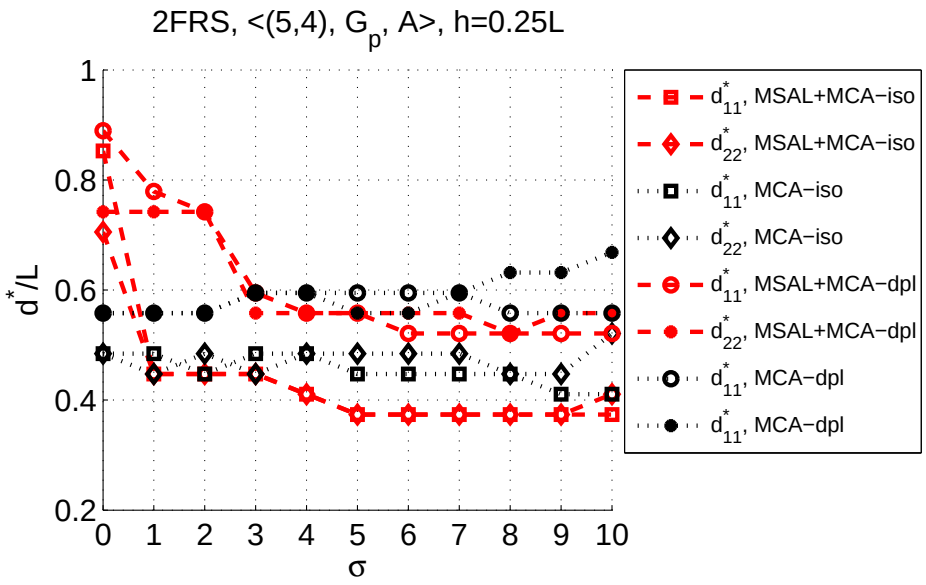

Figure 4.23: Effect of σ on d^*_{ii} with 2FRS

itoring range levels. According to this figure, with both 2FRS and 3FRS, MSAL+MCA performs similar to MSAL at small σ and similar to MCA at large σ values. This is due to the creation of larger number of invalid subareas at high shadowing levels. MSAL seems to be the worst especially under high σ due to its poor availability performance.

Figures 4.29 shows the availability under different channel conditions and floor height with 2FRS. We can observe higher availability with small h and small σ as expected in isotropic case. Small h can provide a larger interval for monitoring ranges to choose from as the best monitoring range, and small σ reflects as fewer invalid subareas. With dipole antennas, the ring shaped adjacent range divides the isotropic range into two parts, which can increase the number of subareas, therefore, the possibility of detecting in invalid subareas. Therefore, as h is increasing, SAL is employed and number of invalid subareas are minimized in order to maximize availability. The effect of invalid subareas on availability is more significant with dipole ranges than with isotropic ranges. 3FRS availability is shown in Figure 4.30. A significant difference is seen with dipole antennas as increasing h causes availability to decrease as opposed to utilizing SAL as in 2FRS.

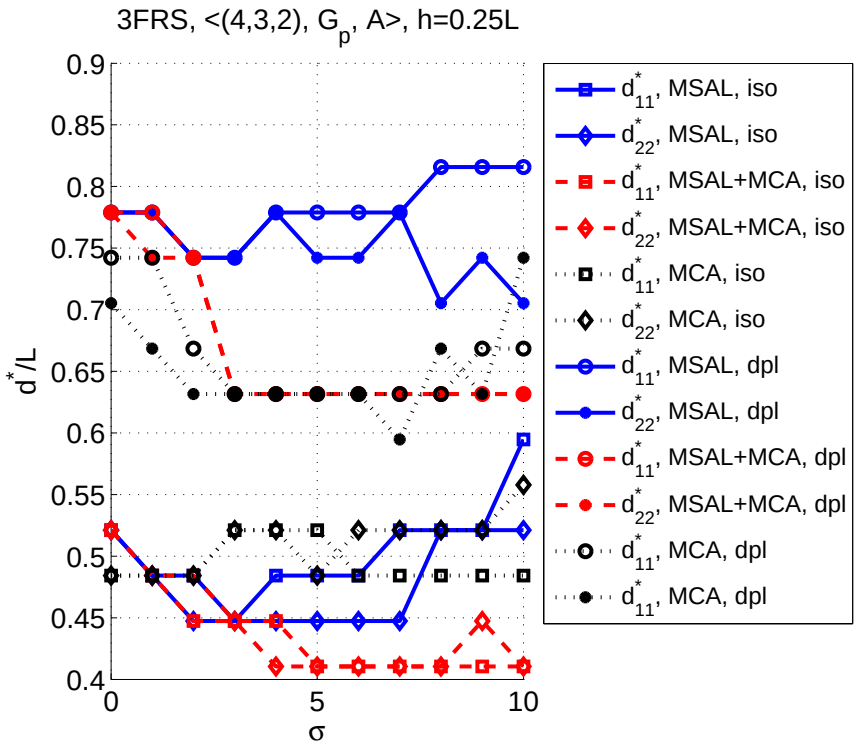

Figure 4.24: Effect of σ on d^*_{ii} with 3FRS

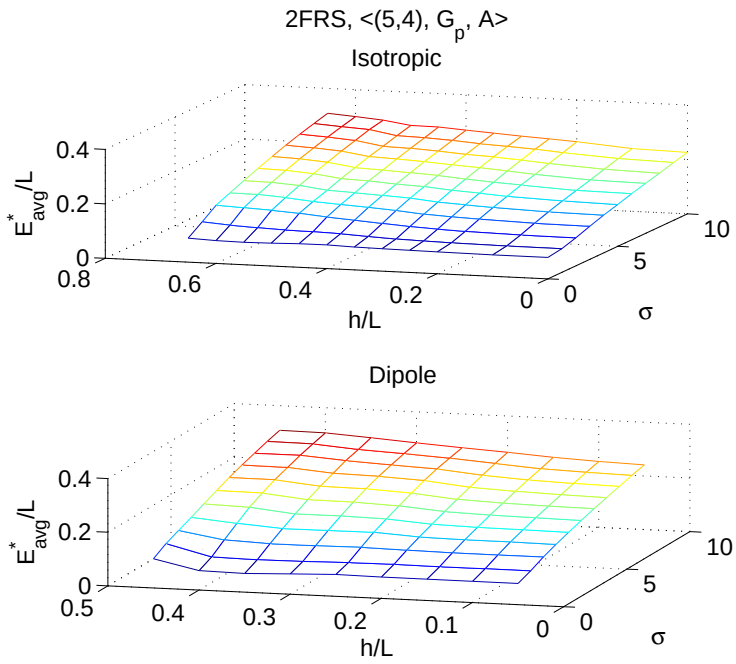

Figure 4.25: E^*_{avg} 2FRS with isotropic and dipole antenna for different values of h and σ

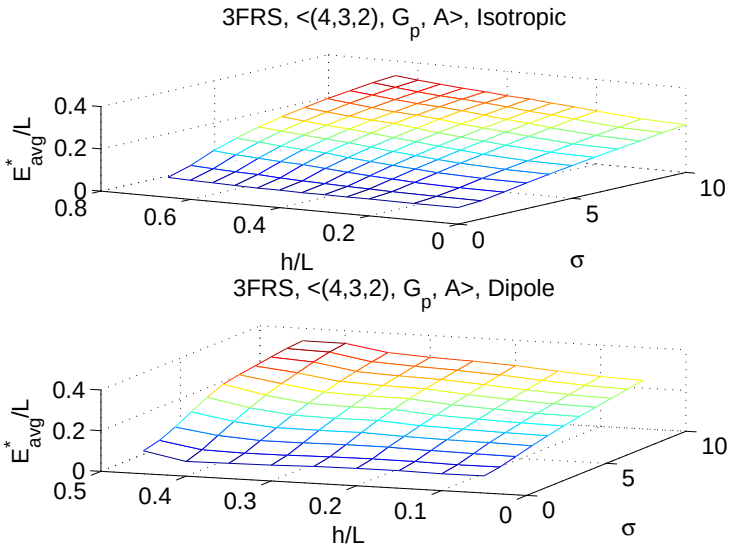

Figure 4.26: E^*_{avg} 3FRS with isotropic and dipole antenna for different values of h and σ

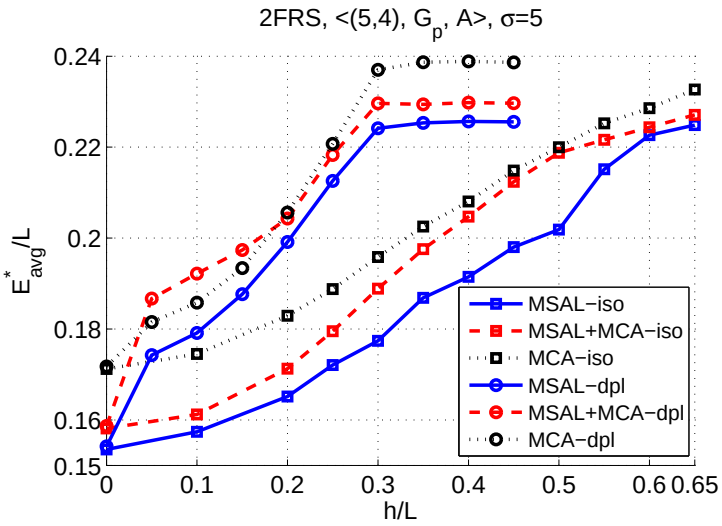

Figure 4.27: E^*_{avg} 2FRS with isotropic and dipole antenna, $\sigma = 5$

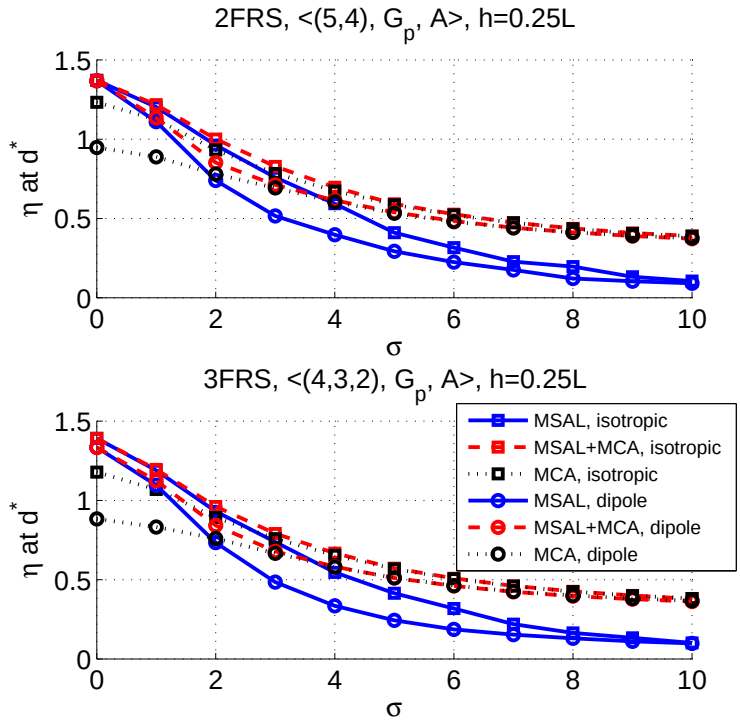

Figure 4.28: Efficiency at d^* with 2FRS and 3FRS, $h = 0.25L$

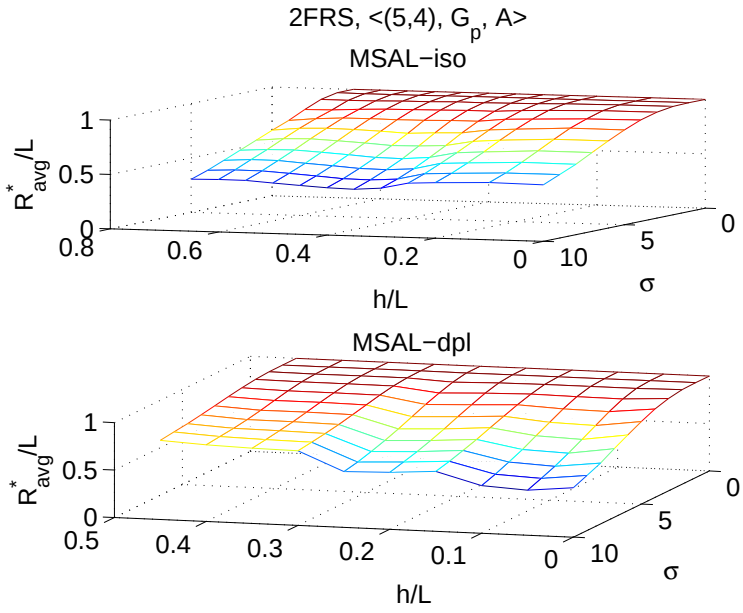

Figure 4.29: R^*_{avg} 2FRS with isotropic and dipole antenna

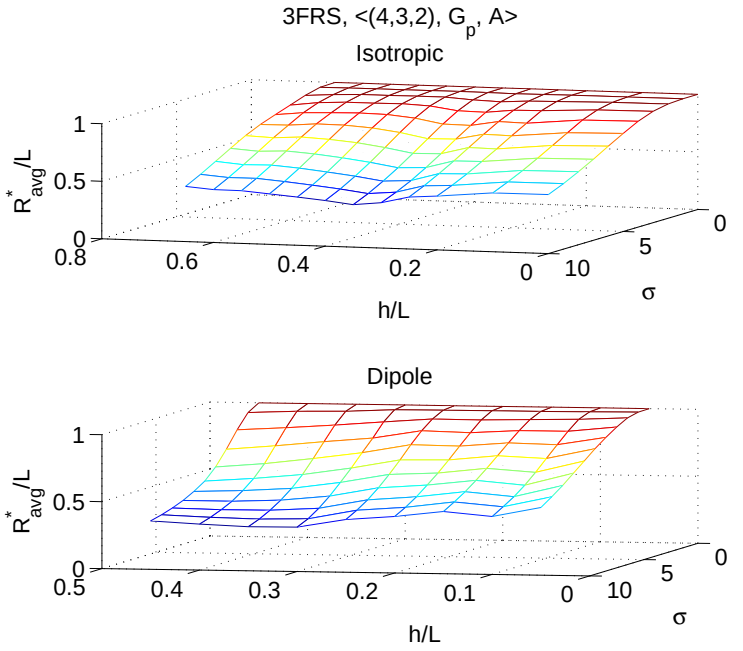

Figure 4.30: R^*_{avg} 3FRS with isotropic and dipole antenna

Figure 4.31 shows the maximum overall efficiency, η^* that can be achieved with 2RFS with respect to h under $\sigma = 0$ and $\sigma = 5$ conditions, and shows that while it is unlikely there will be any control over h, there is a best h value maximizing efficiency when using dipole antennas under ideal channel conditions. The maximum achievable efficiency appears at $h = 0.2L$ for MCA-dpl, and at $h = 0.35L$ for MSAL-dpl when $\sigma = 0$. When $\sigma = 5$, the maximum efficiency is where the floor height is smallest, this is expected as an increase in floor height causes adjacent monitoring ranges to narrow down. A general observation is that MSAL+MCA (same as MSAL when $\sigma = 0$) outperforms other algorithms under any channel conditions. Also, isotropic antenna coverages can provide better efficiency compared to efficiency with dipole antennas. When $h = 0$, SAL is used with $< 9, \mathbf{G}, A >$ scenario, a comparison to multi-floor utilization reveals the advantages of using MoSs across floors over using MoSs on the current floor only. Simulations show that the observations for 2FRS are also valid for 3FRS localization scenario as seen in Figure 4.32, except that the maximum efficiency appears at $h = 0.4L$ for MSAL when $\sigma = 0$. When 2FRS and 3FRS are compared, we see 3FRS efficiency is larger than 2FRS efficiency. This is a reflection of using less number of MoSs per floor on 3FRS scenario. If we zoom into $h = 0.25L$ value to compare performance under different σ, Figure 4.33 shows that MSAL is better than MCA at close to ideal channel conditions, and MSAL+MCA outperforms other algorithms under any channel condition for both 2FRS and 3FRS and for both isotropic and dipole antenna models. Also, localization errors achieved with MSAL+MCA at maximum efficiency levels are the smallest.

4.3.3 Effect of Wall Attenuation Factor

In order to show the effect of wall attenuation factor on MSAL performance we worked with two different floor plans introduced before in Section 3.4.1. In this case, signals traveling from adjacent floors will have to pass through walls to reach MoSs that are placed on the floor of the current floor. We used Eq. 3.37 and 3.38 in order to calculate the received power and detection threshold, and assumed $WAF = 3.1$dB. This attenuation factor causes the MoSs to perceive the range as $d/1.26$. We assume the floor plans and MoS locations are known to the localization system. Detection threshold for each MoS i is calculated separately

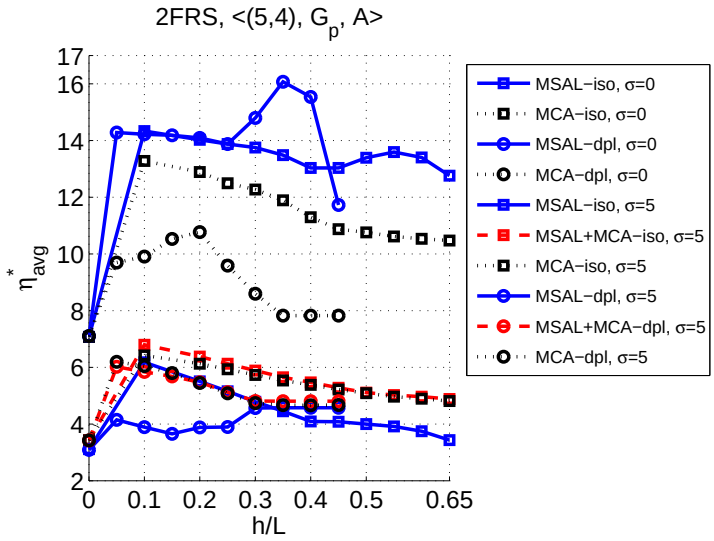

Figure 4.31: Maximum achievable efficiency metric with 2FRS at $\sigma = 0$ and $\sigma = 5$

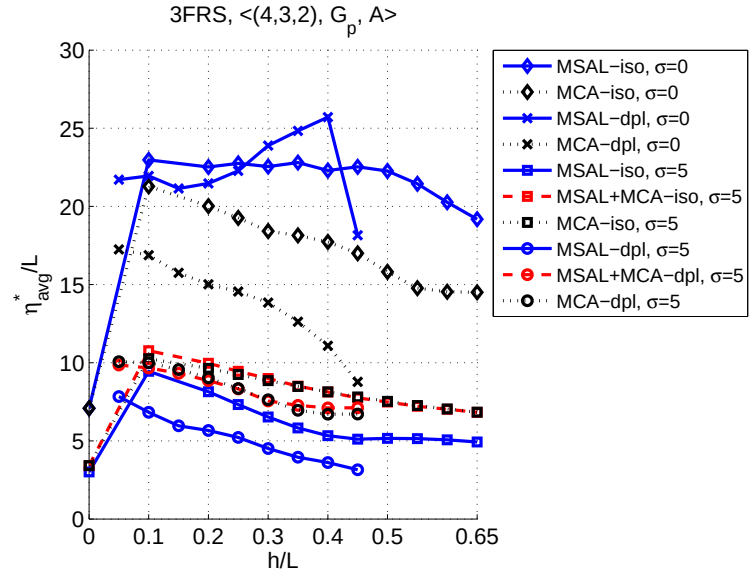

Figure 4.32: Maximum achievable efficiency metric with 3FRS at $\sigma = 0$ and $\sigma = 5$

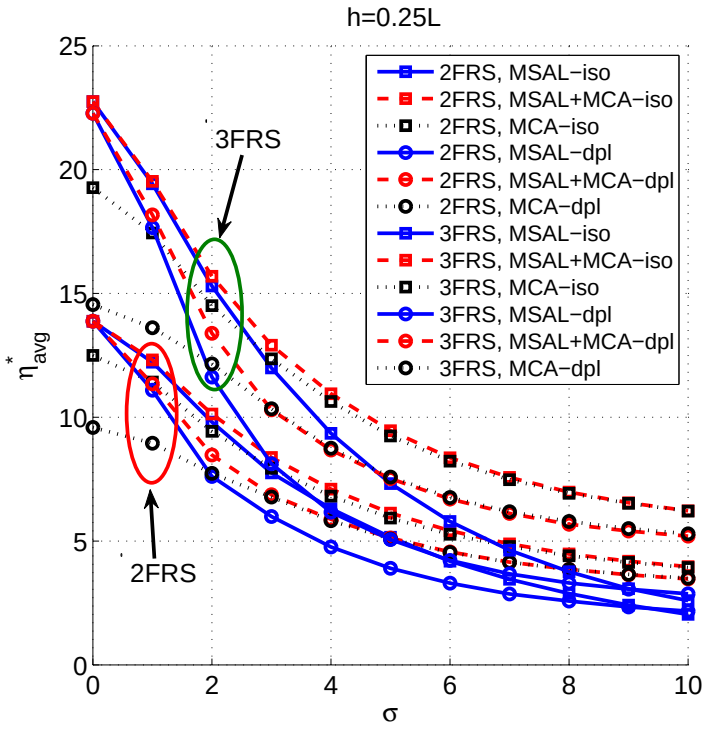

Figure 4.33: Maximum achievable efficiency metric with 2FRS and 3FRS

by considering the average number of walls that a signal will pass through from,

$$R_{th,i} = R_s - n_{avg,i}\, WAF \tag{4.10}$$

In previous sections, R_s is given as,

$$R_s = P_t - 10\alpha log_{10}(d) \tag{4.11}$$

$$R_{th,i} = P_t - 10\alpha log_{10}(d_{MN}) - nWAF \tag{4.12}$$

$$R_s - n_{avg,i}\, WAF = P_t - 10\alpha log_{10}(d_{MN}) - nWAF \tag{4.13}$$

$$P_t - 10\alpha log_{10}(d) - n_{avg,i}\, WAF = P_t - 10\alpha log_{10}(d_{MN}) - nWAF \tag{4.14}$$

$$d_{MN} = \frac{d}{10^{WAF(n-n_{avg,i})/10\alpha}}, \tag{4.15}$$

If $n \geq n_{avg,i}$, MN is detected if only it is in range d, otherwise, it is still detected, and it is detected outside range d as in this case $d_{MN} > d$. Detection threshold for each MoS changes with the average number of walls that a signal must pass through to reach this MoS. But if the signal actually passes through a less number of walls than the average, this threshold allows MNs that are outside of monitoring range to be detected as if they are in the range. Due to this condition, invalid subareas can appear even under ideal channel conditions. These MNs can be located by MSAL+MCA or MCA algorithm. Therefore, we expect to see different accuracy for MSAL and MSAL+MCA under ideal conditions.

WAF effect is investigated using 2 localization scenarios, 2FRS and 3FRS. Figures 4.34 and 4.35 show how accuracy changes with floor height in the presence of wall attenuation factor assuming no shadowing with 2FRS and 3FRS localization scenarios, respectively. We observe that the best accuracy is achieved at smaller floor height due to larger monitoring ranges. Performance of three algorithms are similar with isotropic antennas, whereas with dipole antennas, MSAL outperforms the other two algorithms. Previously, we showed that MSAL and MSAL+MCA performs exactly the same due to no invalid subareas under ideal channel conditions. However, when WAF is considered, invalid subareas may be formed due to the ambiguity of the number of walls that a signal may pass through to reach MoSs. Then, MSAL is not expected to localize the nodes that fall into the invalid subareas; but MSAL+MCA is expected. In these two figures, MSAL+MCA has worse accuracy than

MSAL accuracy because MSAL+MCA average error includes the localization error of nodes in invalid subareas.

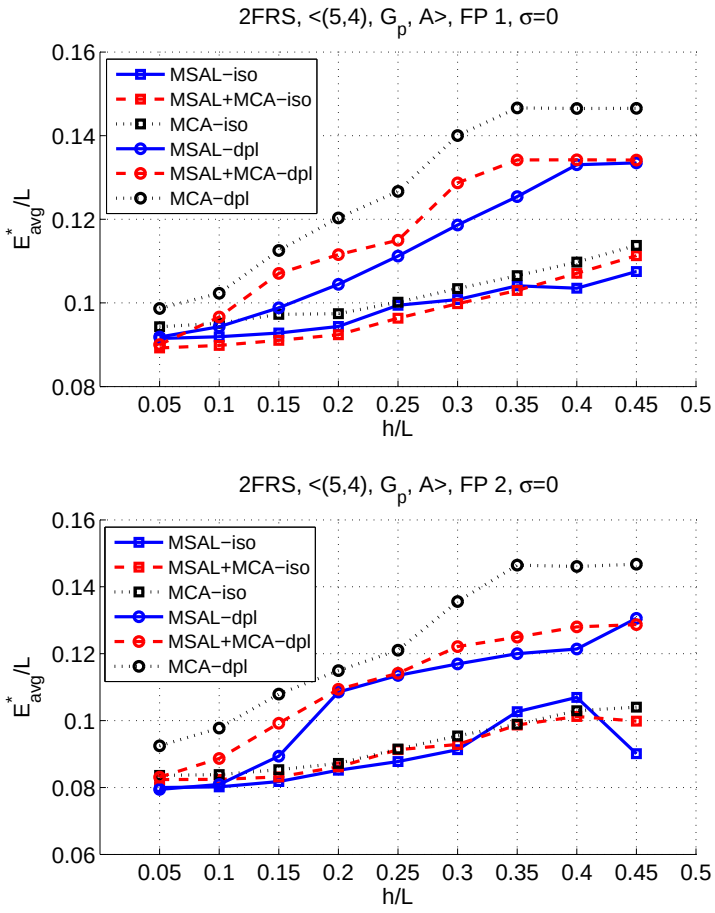

Figure 4.34: Accuracy versus h at $\sigma = 0$ with 2FRS

In Figure 4.36, we demonstrate an instance of how MSAL+MCA localizes MNs shown with blue stars in 2 floors in floor plan 1. Solid lines show the coverage with R_s and, dashed lines show the coverage with $R_{th,i}$. Red diamonds show the potential estimated locations with MSAL used to localize MNs detected in valid subareas, whereas green diamonds show the potential estimated locations with MCA used to localize nodes detected in invalid subareas. Note that with MSAL, even MNs out of the range with $R_{th,i}$ can be located as if they are in range. This is due to relationship between n and $n_{avg,i}$.

Figures 4.37 and 4.38 show the availability with 2FRS and 3FRS. These figures confirm

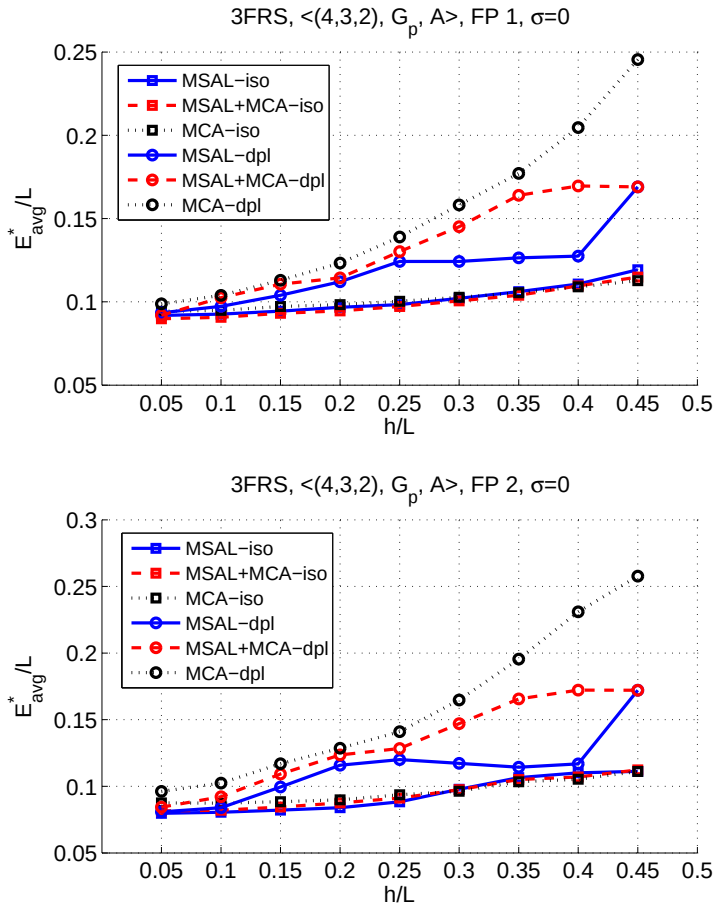

Figure 4.35: Minimum error versus h at $\sigma = 0$ with 3FRS

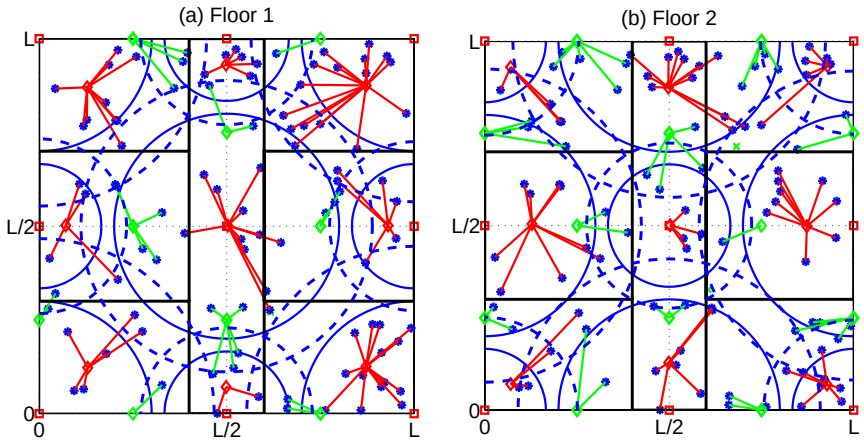

Figure 4.36: Visualization of estimated locations with 2FRS with Floor Plan 1, $h = 0.25L$

the discussion about the coverage and invalid subareas with MSAL being lower than 100% even under ideal channel conditions. MSAL+MCA and MCA availabilities are less affected by WAF when compared to MSAL. Also, we note that although MSAL cannot have high availability performance, it can contribute to MSAL+MCA with reduced localization error for an average of $100R_{MSAL}\%$ of the floor area.

Figure 4.39 shows the ratio of best monitoring ranges in case of no shadowing. Although ratio of monitoring ranges stays close to unity, it changes with different floor height.

The good accuracy/availability performance of MSAL+MCA reflected as good performance in terms of the overall efficiency in both localization scenarios and both floor plans. Simulation results are shown in Figures 4.40 and 4.41 with 2FRS and 3FRS localization scenarios. Isotropic antennas is also shown to be better in terms of the overall localization efficiency. 3FRS can provide better efficiency than 2FRS especially when floor height is small. When floor height is small, efficiency levels with isotropic and dipole antennas area similar due to smaller difference in vertical range. As floor height increases, the efficiency levels with isotropic and dipole antennas diverge from each other. Dipole antennas can provide less efficiency as expected. For example, for FP 1, efficiency levels with 3FRS range

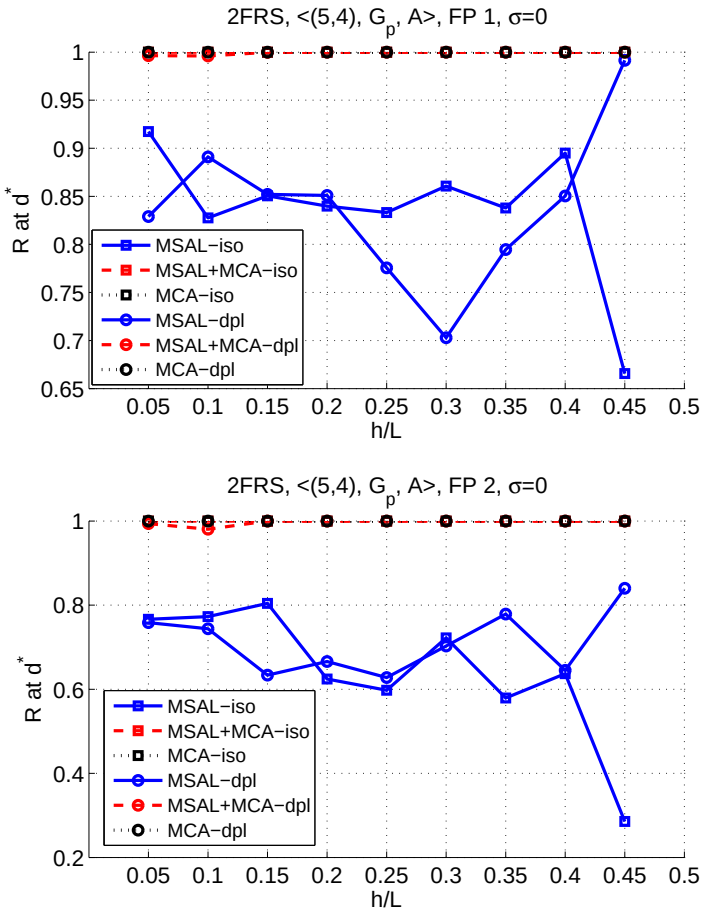

Figure 4.37: Availability at minimum error conditions versus h at $\sigma = 0$ with 2FRS

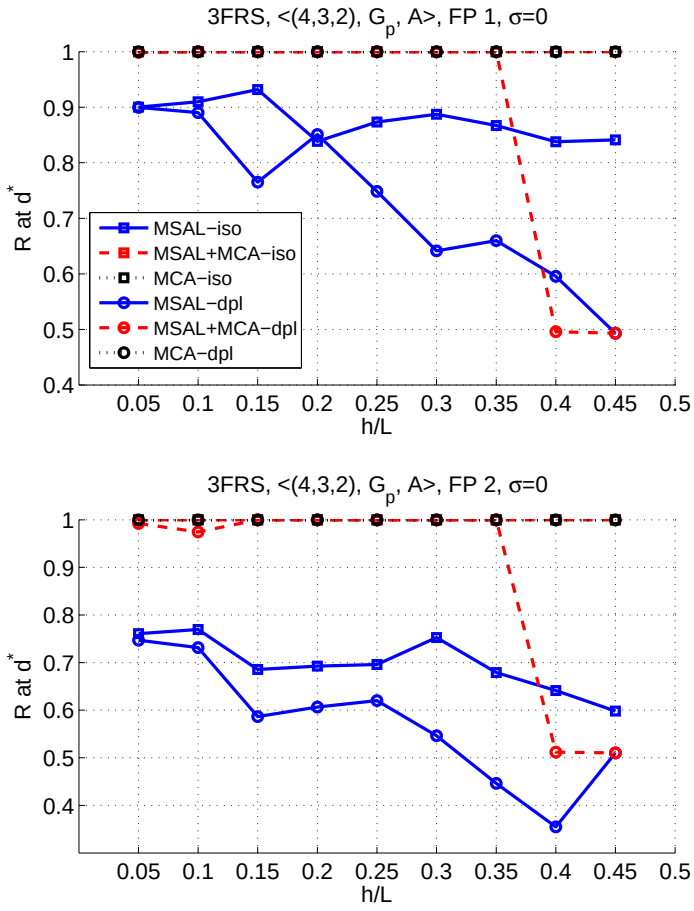

Figure 4.38: Availability at minimum error conditions versus h at $\sigma = 0$ with 3FRS

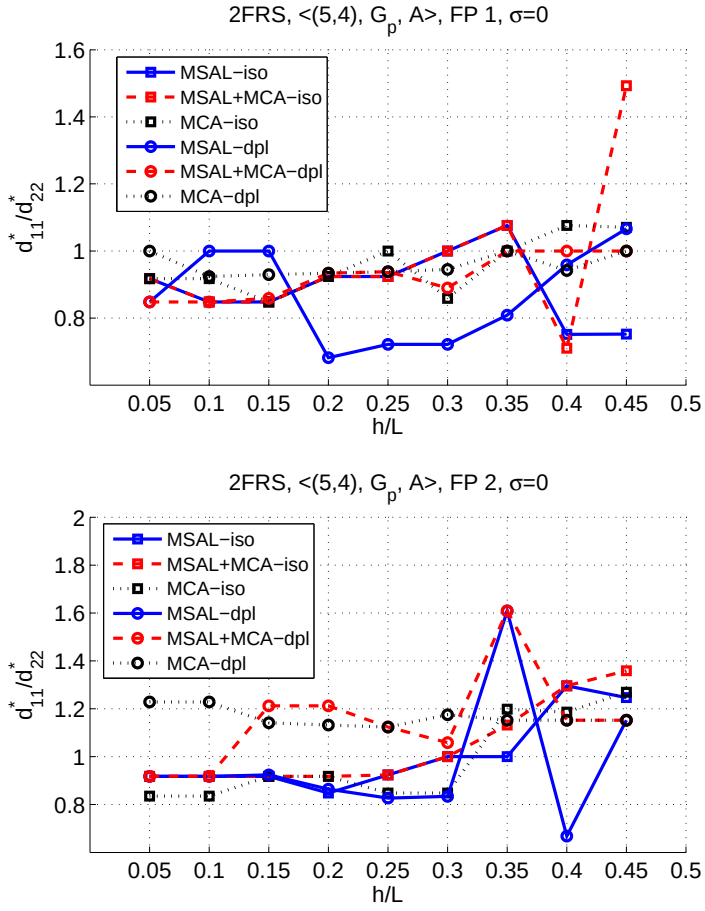

Figure 4.39: Minimum error monitoring range versus h at $\sigma = 0$ with 2FRS

from 20 to 16 with isotropic antennas and 20 to 10 with dipole antennas, whereas with 2FRS levels range from 13 to 10 with isotropic antennas and 13 to 8 with dipole antennas. When we compare best efficiency levels at $h = 0.25L$ with and without WAF effect, we observe 17% degradation in efficiency with 3FRS and 14% degradation with 2FRS using MSAL-iso. MSAL-dpl provides even worse efficiency.

We ran a set of simulations to show the effect of WAF while the channel is subject to log normal shadowing. Figures 4.42 and 4.43 show the accuracy with respect to h under shadowing with $\sigma = 5$ with 2FRS and 3FRS, respectively. Similar results such as MSAL outperforms other algorithms, and 2FRS can provide better accuracy can be drawn from these figures. In addition to these, we observe that with dipole antennas, when $h \leq 0.3$, MSAL and MSAL+MCA approach to similar accuracy levels.

As expected, large floor height and large σ can degrade the maximum efficiency of localization system. Figures 4.44 and 4.45 show the overall efficiency with 2FRS and 3FRS with respect to σ when $h = 0.25L$. In both cases, increasing shadowing effect degrades overall efficiency, and MSAL+MCA outperforms the other algorithms. 3FRS has higher efficiency than 2FRS due to smaller number of MoSs used per floor. Also, at high levels of shadowing, MSAL-dpl outperforms MSAL-iso due to a higher number of invalid subareas created by a larger adjacent monitoring range with isotropic antennas.

4.3.4 Effect of Floor Attenuation Factor

In multi-floor buildings, a signal propagation model including floor attenuation factor may be used to simulate the signal power degradation when it travels from another floor [35]. In this case, received power is attenuated by FAF dB as in,

$$P_r = P_t - 10\alpha log(d_{ij}) + X - FAF(f) \tag{4.16}$$

where FAF represents the floor attenuation factor which is a function of number of floors that a signal pass through to reach the receiver. However, the attenuation is not a linear function. FAF for one floor was calculated empirically at 914Mz in [35] as 13dB. 13dB attenuation on received power cuts the traveling distance of this signal by $x_f = 10^{FAF/(10\alpha)} \approx 2.7$. This

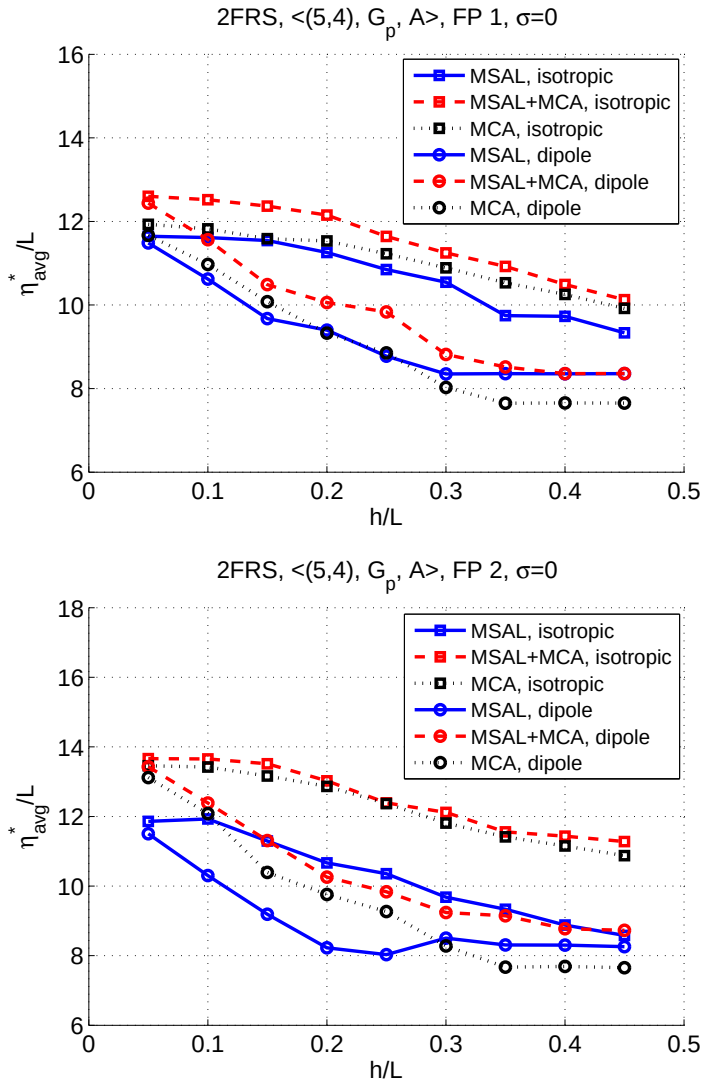

Figure 4.40: Maximum efficiency versus h at $\sigma = 0$ with 2FRS

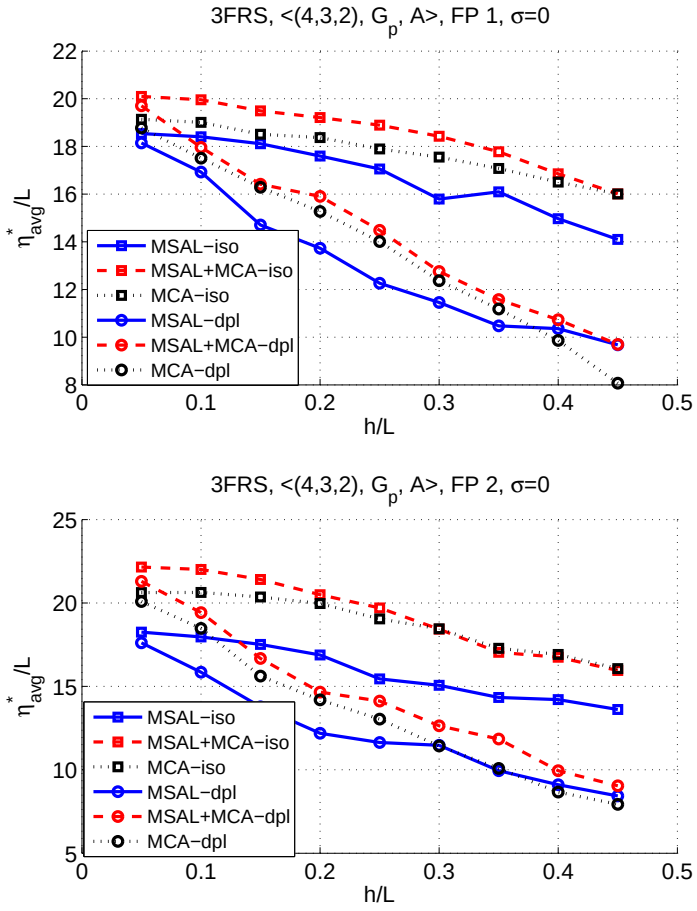

Figure 4.41: Maximum efficiency versus h at $\sigma = 0$ with 3FRS

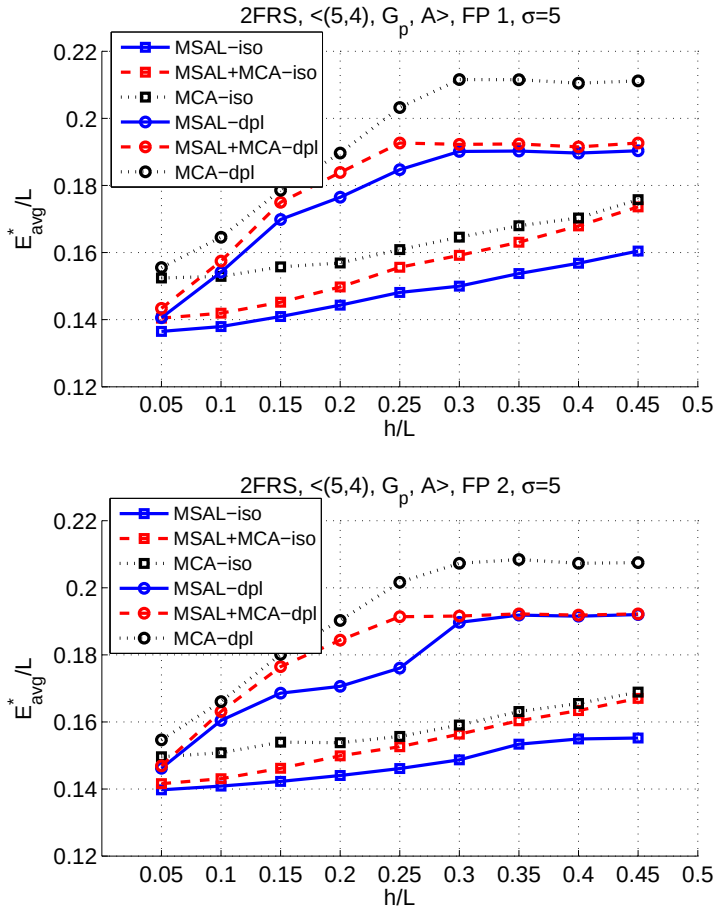

Figure 4.42: Minimum error for different values of h at $\sigma = 5$ with 2FRS

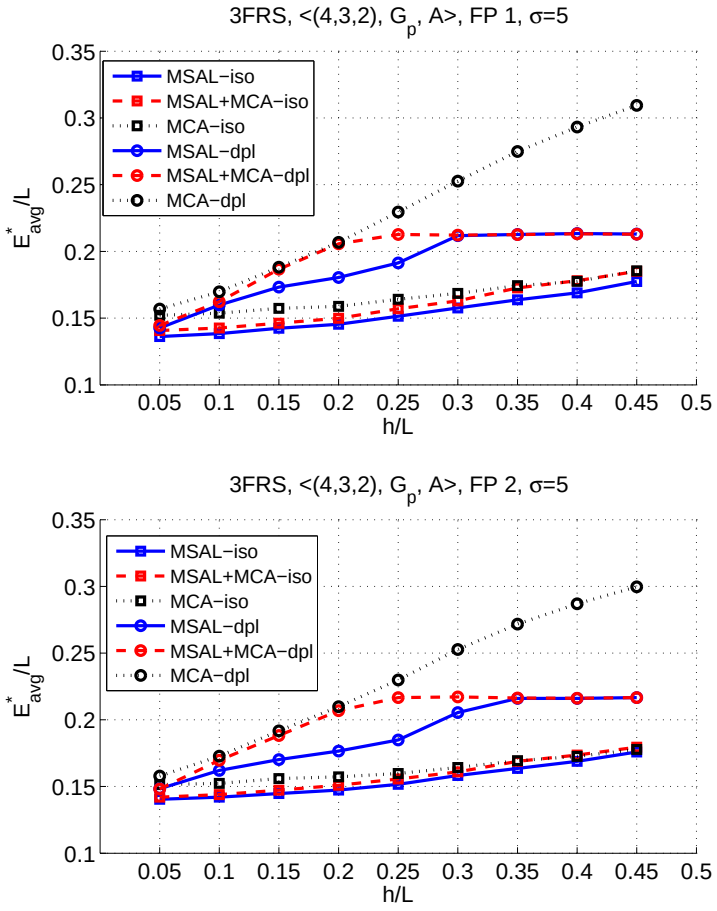

Figure 4.43: Minimum error for different values of h at $\sigma = 5$ with 3FRS

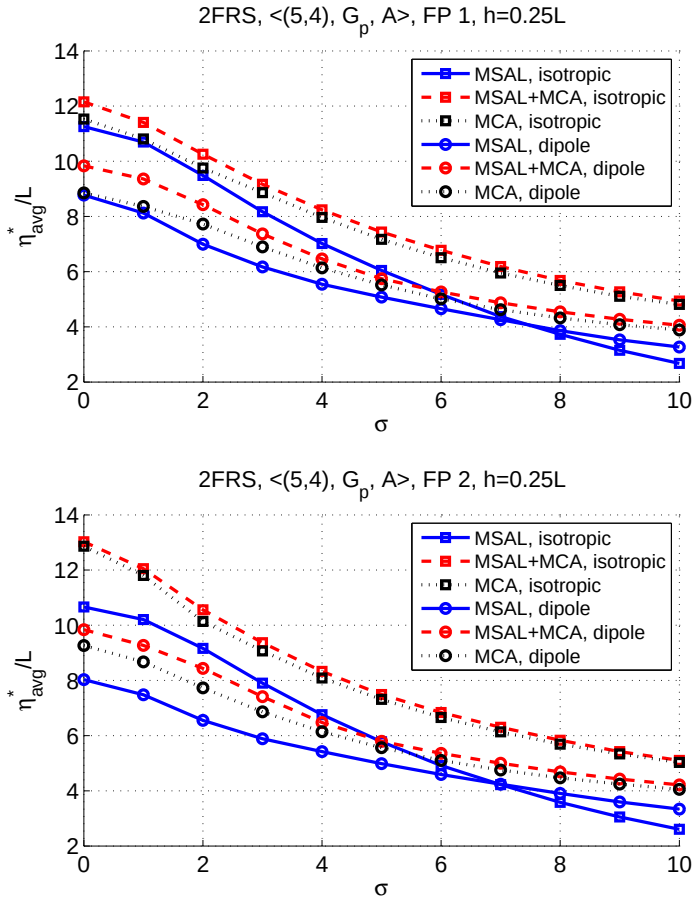

Figure 4.44: Maximum efficiency for different values of σ and $h = 0.25L$ with 2FRS

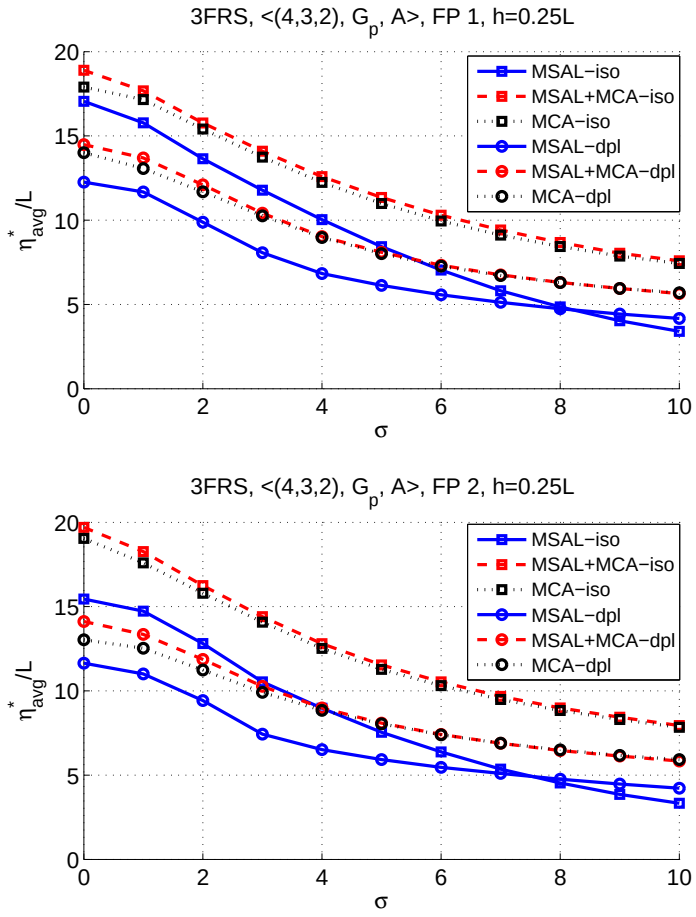

Figure 4.45: Maximum efficiency for different values of σ and $h = 0.25L$ with 3FRS

case is equivalent to increasing the floor height by a factor of x_f, i.e., $h_{FAF} = h \times x_f$. The effective floor height is shown in Figure 4.46. From the relationship between h and h_{FAF}, required monitoring range for multi-floor coverage and allowed current floor range ($d_{ii} \leq L$), we expect to see single floor localization is used with $h_{FAF} > L$ with isotropic antennas and $2h_{FAF} > L$ with dipole antennas instead of multi-floor localization. then, the floor height thresholds for multi-floor localization with isotropic antenna antenna and dipole antennas are $h > \frac{L}{x_f}$ and $h > \frac{L}{2x_f}$, respectively.

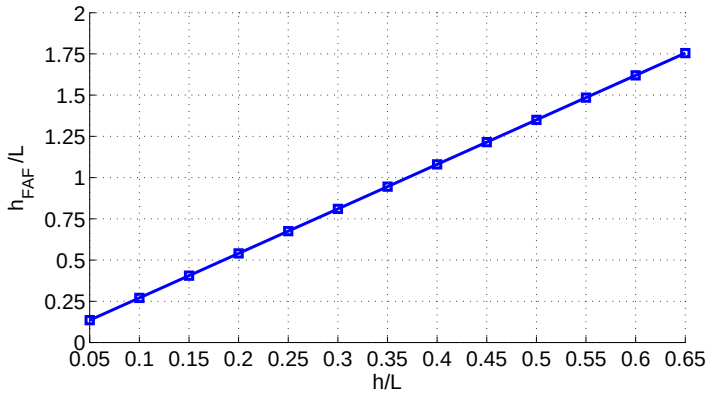

Figure 4.46: Effective floor height when $FAF = 13$dB

FAF effect is investigated using two localization scenarios, 2FRS and 3FRS, as in previous sections. Figures 4.47 and 4.48 show how accuracy changes with floor height in the presence of floor attenuation factor assuming no shadowing. From this figure, we observe that minimum localization error with dipole and isotropic antennas increases with increasing floor height, and saturates at the same value for floor height larger than $0.35L$. We also see a jump in accuracy of MSAL-dpl when $h \leq 0.15L$. Due to floor attenuation effect, MSAL-dpl and MCA-dpl use MoSs only on the current floor to get the best accuracy when $h \geq 0.2L$ and $h \geq 0.15L$, respectively. Isotropic antennas have a wider adjacent floor range, and the constraint on floor size is more loose ($d > h$) than with dipole antennas ($d > 2h$). This is why MSAL-iso and MCA-iso use MoSs only on the current floor to get the best accuracy when $h \geq 0.4L$ and $h \geq 0.35L$, respectively. Accuracy with isotropic antennas and dipole antennas approaches to same values as the range of dipole antennas will be equivalent to range of isotropic antennas on the current floor.

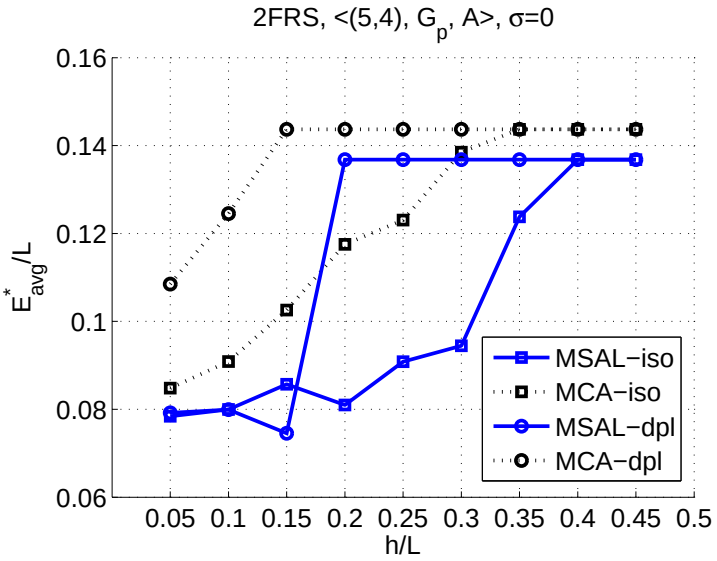

Figure 4.47: Minimum error for different values of h at $\sigma = 0$ with 2FRS

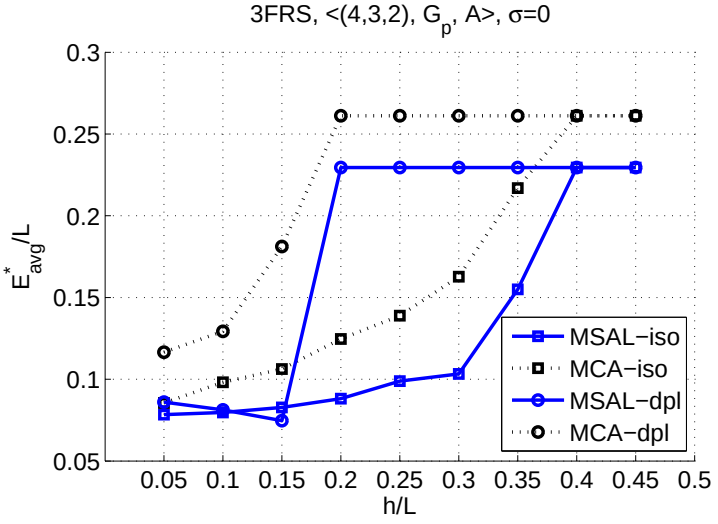

Figure 4.48: Minimum error for different values of h at $\sigma = 0$ with 3FRS

Figures 4.49 and 4.50 show the ratio of best monitoring ranges in case of no shadowing.

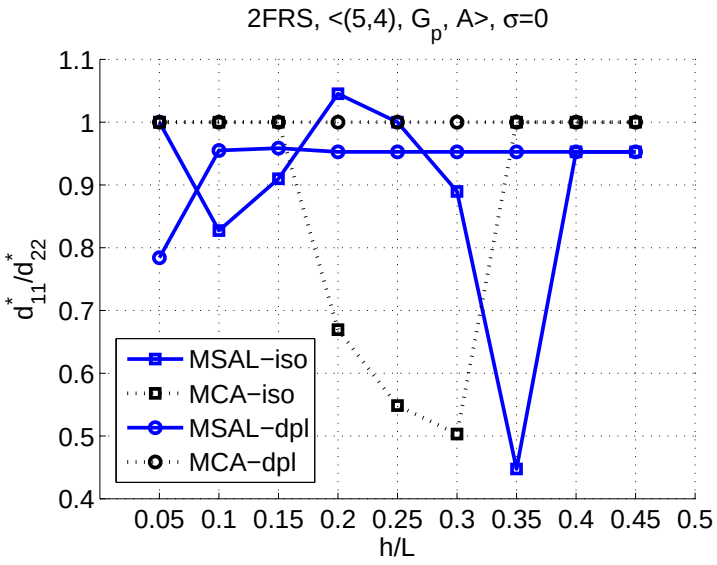

Figure 4.49: Minimum error monitoring range for different values of h at $\sigma = 0$ with 2FRS

Figures 4.51 and 4.52 show the accuracy with 2FRS and 3FRS under shadowing with $\sigma = 5$, respectively. Also under shadowing we observe the similar behavior as in ideal channel conditions. However, in this case, accuracy with isotropic and dipole antennas approach to the same value at a lower floor height than with ideal channel. Availability under ideal channel conditions is equal to unity; however, as shown in Figures 4.53 and 4.54, the availability under $X = N(0, 5)$ decreases slightly with increasing height in case of MCA and MSAL+MCA. In case of MSAL, availability first decreases and then starts to increase with increasing floor height. While using MoSs across floors as floor height increases availability decreases due to both coverage and effect of invalid subareas created by largest monitoring ranges that are preferred in order to utilize multi-floor localization. For example with 2FRS, when $h \geq 0.25L$, monitoring range stays at a level that it cannot utilize multi-floor localization if under ideal conditions. However, with shadowing, there are still areas that can be heard by MoSs on multiple floors. As floor height increases these areas gets smaller and eventually disappears. This is the reason why availability starts to increase and approach the same values as with dipole antennas while h is increasing. At $h = 0.15L$, MSAL-dpl uses MoSs on current floor only, therefore the availability increases due to less

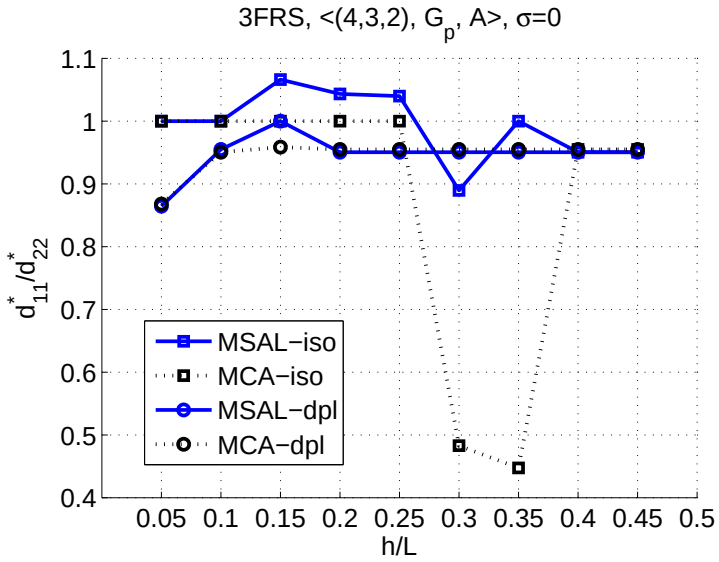

Figure 4.50: Minimum error monitoring range for different values of h at $\sigma = 0$ with 3FRS

number of MoSs, and smaller probability of invalid subareas.

Figures 4.55 and 4.56 show the ratio of best monitoring ranges in case of shadowing with $\sigma = 5$.

Figures 4.57 and 4.58 show how the accuracy is affected by shadowing under different floor height assumptions in the presence of FAF with 2FRS. The observation is that for floor smaller floor height the advantage of MSAL and MSAL+MCA is obvious when compared to MCA even at high levels of shadowing. This advantage starts to disappear at larger floor height values. Figures 4.59 and 4.60 show how the accuracy is affected by shadowing with 3FRS. The outcome of this simulation is similar to the outcome from 2FRS simulations.

4.3.5 Effect of Random Placement

It may not be possible or feasible to place MoSs on a grid on each floor of a building. Therefore, in this section we provide the results of a simulation of 2FRS localization scenario with randomly placed MoSs, $< (5,4), \mathbf{R}, A >$. In order to make a fair comparison between grid placement scenarios and random placement scenario, we used the same number of MoSs

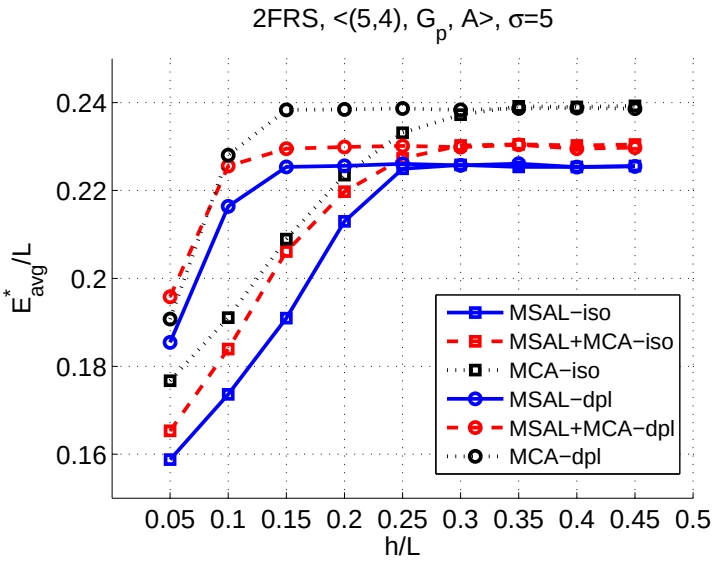

Figure 4.51: Minimum error for different values of h at $\sigma = 5$ with 2FRS

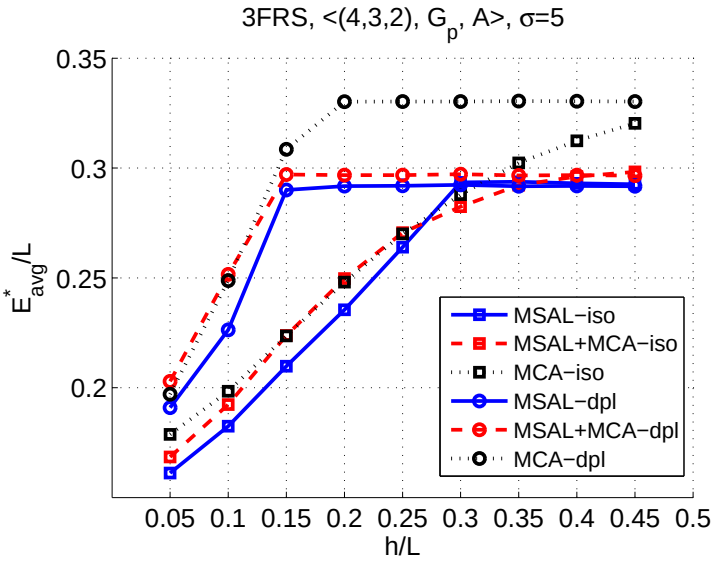

Figure 4.52: Minimum error for different values of h at $\sigma = 5$ with 3FRS

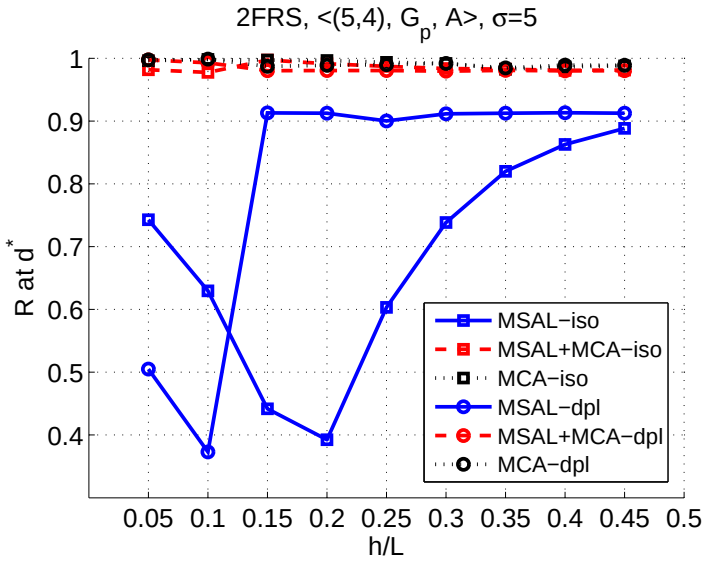

Figure 4.53: Availability at minimum error conditions for different values of h at $\sigma = 5$ with 2FRS

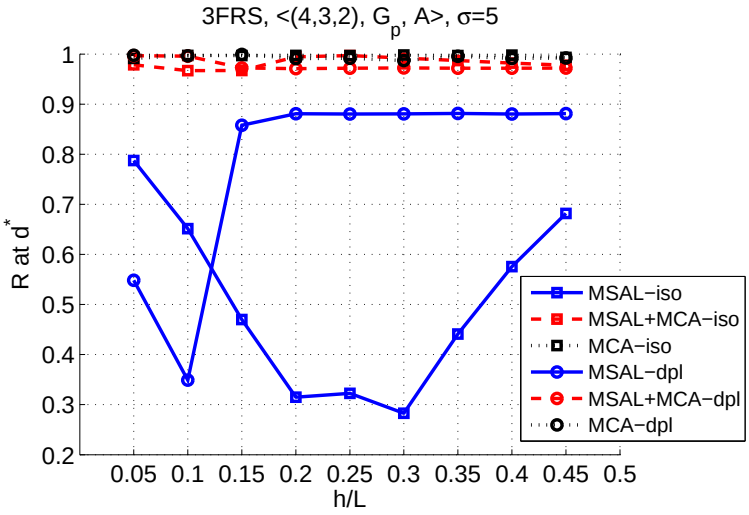

Figure 4.54: Availability at minimum error conditions for different values of h at $\sigma = 5$ with 3FRS

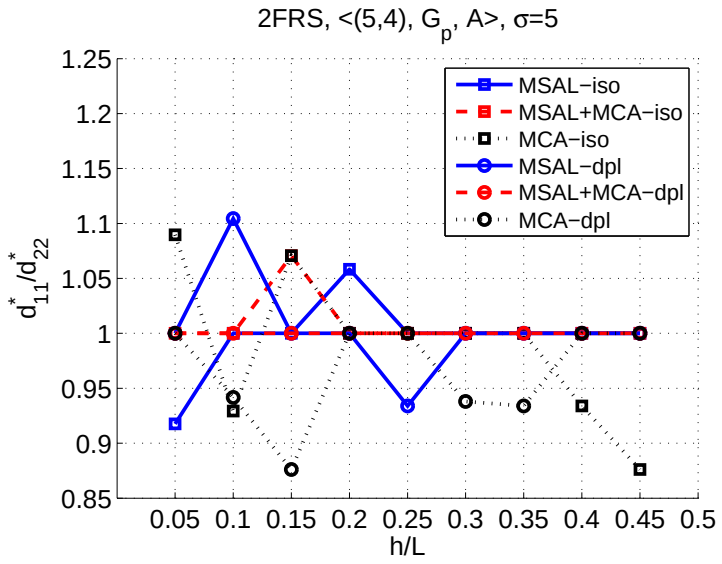

Figure 4.55: Minimum error monitoring range for different values of h at $\sigma = 5$ with 2FRS

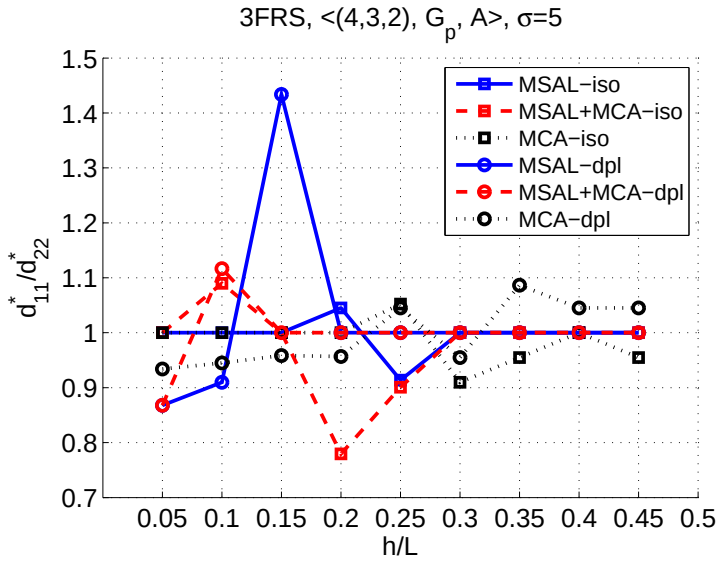

Figure 4.56: Minimum error monitoring range for different values of h at $\sigma = 5$ with 3FRS

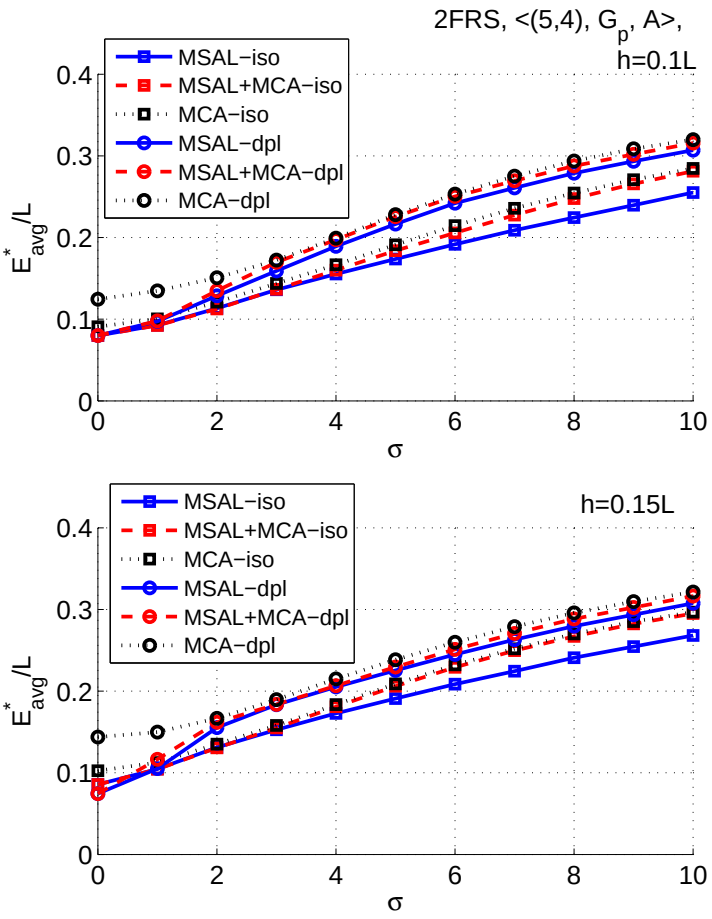

Figure 4.57: Minimum error monitoring range for different values of s with 2FRS

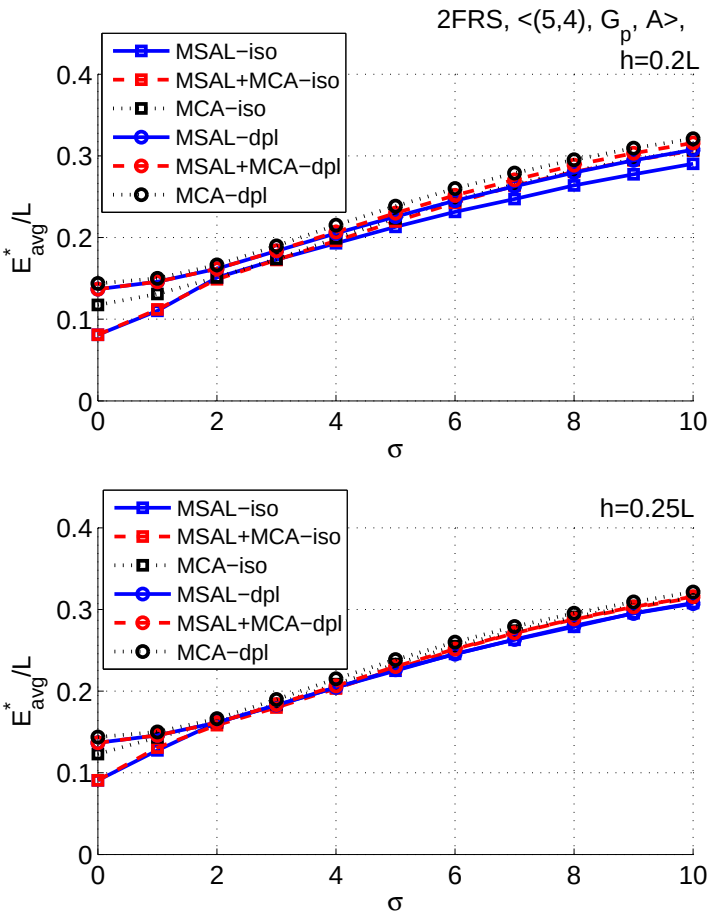

Figure 4.58: Minimum error monitoring range for different values of s with 2FRS

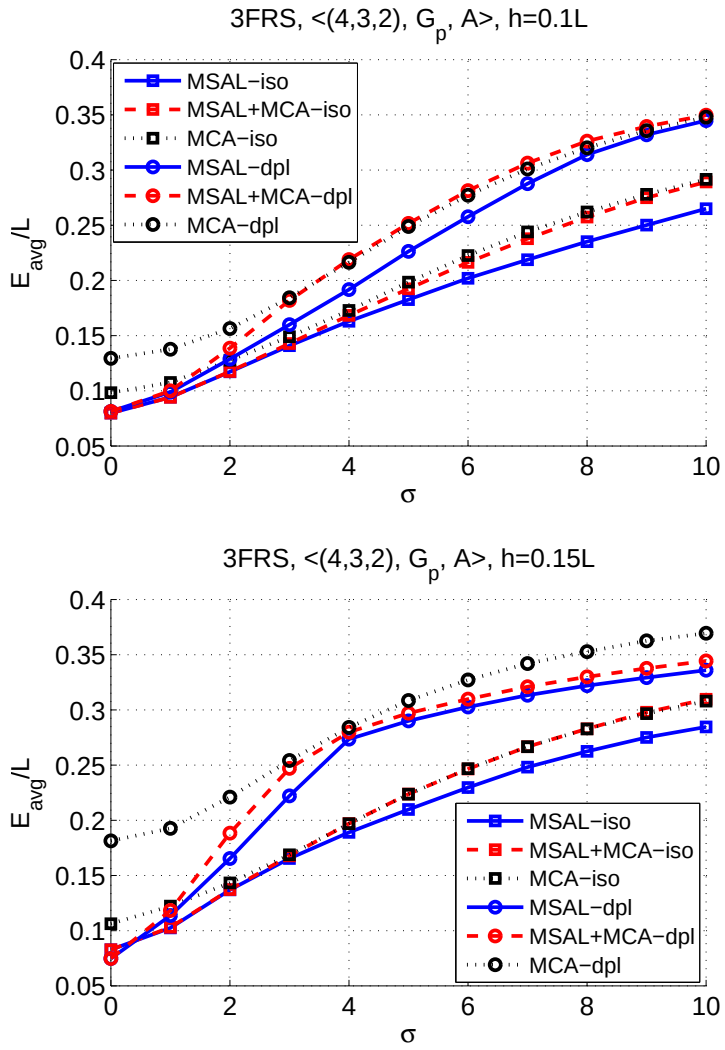

Figure 4.59: Minimum error monitoring range for different values of s with 3FRS

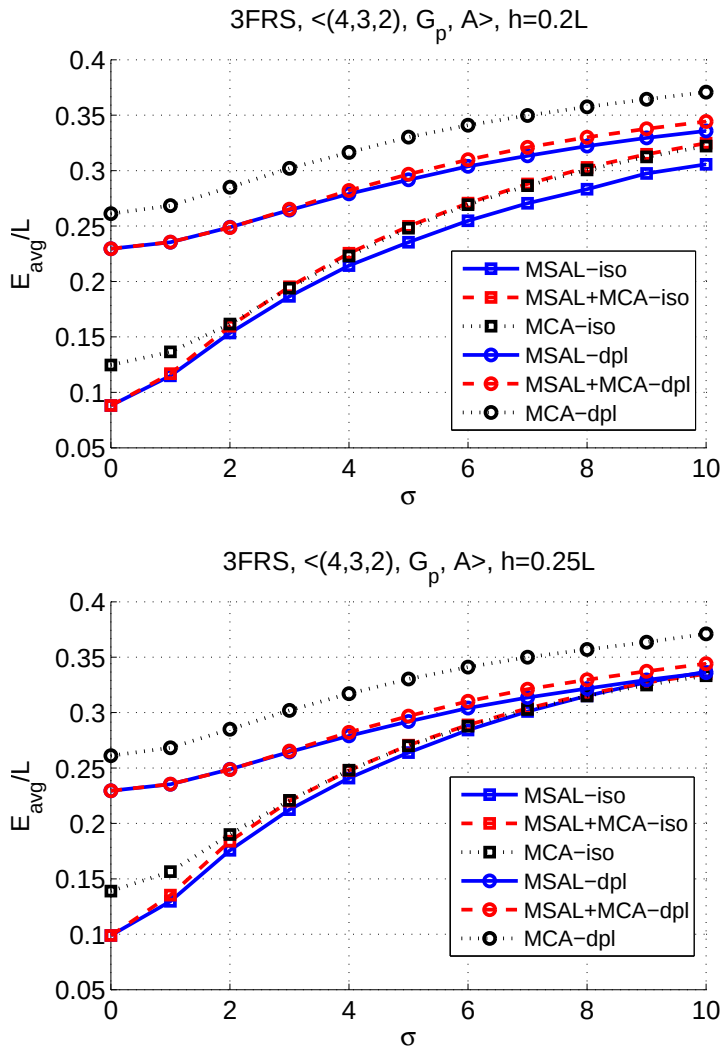

Figure 4.60: Minimum error monitoring range for different values of s with 3FRS

in two floors as 5 and 4, 9 in total. At each floor 4 MoSs are randomly placed at 4 small divisions of each floor, and every other floor one MoS is placed at the center of the floor. Simulation results shown are the average of 20 such placement of MoSs in 2 floors.

The accuracy of random placement scenario on two floors are $0.0873L$ and $0.0879L$, and average accuracy on both floors is $0.0876L$ achieved at $(d_{11}, d_{22}) = (0.48L, 0.62L)$. These monitoring ranges can provide 100% availability on the average. The variance between the performance of different placement instances are shown in Figure 4.61: the variance is very small and the worst accuracy achieved is $0.024L$ larger than the best accuracy.

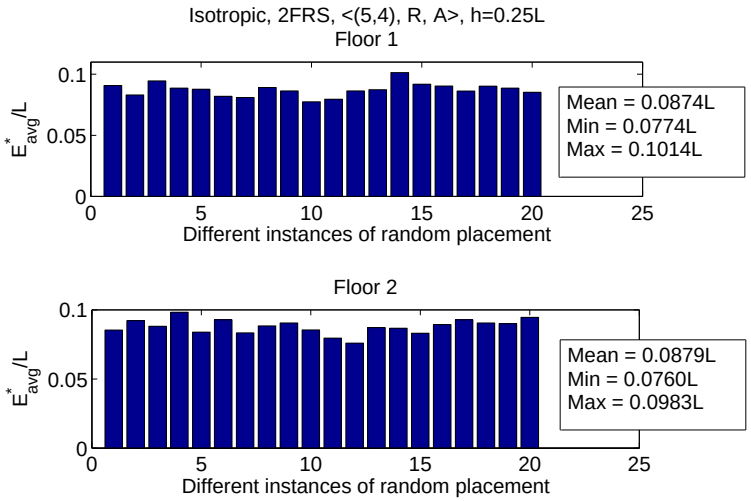

Figure 4.61: Accuracy of different instances of random placement with 2FRS, $< (5,4), \mathbf{R}, A >$, $\sigma = 0$

Figures 4.62 and 4.63 show the minimum localization error and efficiency at minimum error conditions under different floor height assumptions and under different channel conditions. According to this, under ideal conditions, accuracy achieved with MSAL is similar to the one achieved with scenario $< (5,4), \mathbf{G}_p, A >$. In terms of efficiency MSAL+MCA outperforms other algorithms under shadowing, and MSAL performs better than MCA under ideal conditions due to better accuracy and same 100% availability.

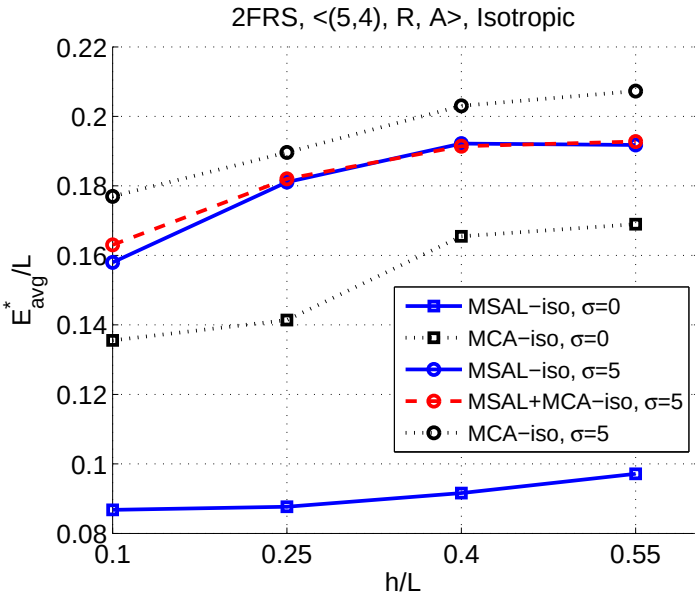

Figure 4.62: Accuracy with 2FRS, $< (5,4), \mathbf{R}, A >$

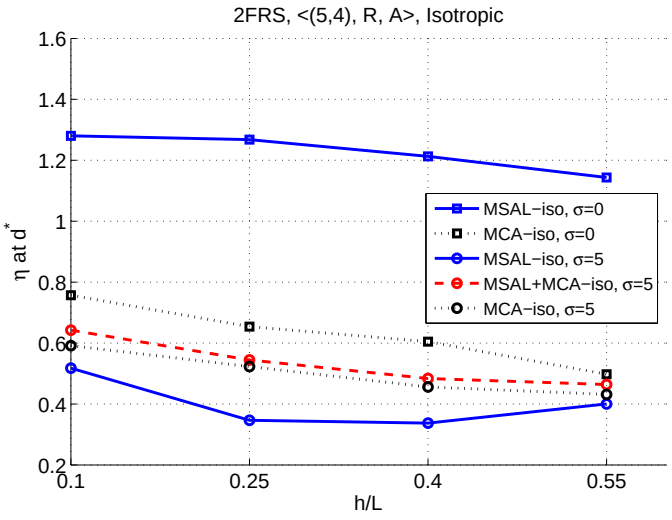

Figure 4.63: Efficiency with 2FRS, $< (5,4), \mathbf{R}, A >$

4.4 FLOOR DETERMINATION FOR MSAL

In Section 4.1, we introduced the multi-floor subarea algorithm and in the following sections, we demonstrated that in multi-floor buildings, the monitoring stations across floors can be utilized in order to improve the efficiency of localization system by keeping the number of infrastructure elements minimal without increasing the localization error. Our objective was to show that employing MoSs across floors is advantageous in terms of cost and accuracy at the same time. For this study, we assumed that an external floor determination (FD) algorithm was available to provide the floor information of mobile nodes.

In this section, we propose that a majority based floor determination algorithm can be used to find which floor a MN is residing in. When floor information is not correct, it is not worthwhile to find the two dimensional location of the MN on the wrongly estimated floor. Therefore, the availability of the localization algorithm will be affected by the FD algorithm's performance. The FD performance on the other hand will get affected by the number and placement of MoSs on adjacent floors, and also the antenna type. For FD, the monitoring range on the vertical direction is very important, because inputs from multiple floors are needed in order to decide which floor the MN is located at.

In the following, we first explain how majority based floor determination algorithm works, then we analyze the effect of antenna type and its relationship with the floor height. Then, we investigate the floor determination performance under log normal shadowing. And we continue with the evaluation of multi floor localization performance when integrated with floor determination algorithm.

4.4.1 Majority Based Floor Determination (MBFD)

Exploiting MoSs on multiple floors to improve efficiency implies that it is now necessary to use the RSS observed at MoSs to determine the correct floor on which a MN is located. We determine the floor based on the sum of m strongest RSS values observed at the MoSs on each floor that detect the same MN. This algorithm is called majority based floor determination (MBFD) [46].

Let the sum of the m strongest RSSs observed in set MoS_i be $RSS^{(m)}_{sum,i}$, where MoS_i is the set of MoSs in floor i that detect the same MN. We call m the decision depth. Then the algorithm decides that the MN is located on floor with the highest sum of m RSS measurements:

$$f^* = argmax_{f=1,\ldots,F}\left(RSS^{(m)}_{sum,f}\right) \tag{4.17}$$

4.4.1.1 Special case When $m = 1$, this algorithm only compares the strongest RSSs observed in different floors.

In ideal conditions (no shadowing), for this algorithm with $m = 1$ to estimate the floor of a MN correctly as floor i, following two constraints have to be satisfied:

1. The first constraint is that the monitoring range d_{ii} of MoSs in MoS_i must be sufficiently large so that the area A of floor i is covered fully.
2. The second constraint exists due to the relationship between the placement of MoSs, the floor height, and antenna type: Let MN^*_i be a MN on floor i that is located at the farthest physical distance from any MoSs $\in \text{MoS}_i$. And let the MoS^*_i is the closest MoS to MN^*_i on floor i,$MoS^*_i = argmin_k\left(dist[MN^*_i, MoS_{ik}]\right)$. Then, the distance MN^*_i to its closest MoS, $dist[MN^*_i, MoS^*_i]$, on floor i must be upper bounded by the minimum distance of MN^*_i to a MoS on floor j that can detect it, $min_k\left(dist[MN^*_i, MoS_{jk}]\right)$, so that MN^*_i is closer to a MoS on the same floor and not one on a different floor (j):

$$min_k\left(dist[MN^*_i, MoS_{ik}]\right) < min_k\left(dist[MN^*_i, MoS_{jk}]\right),\ j \neq i, \tag{4.18}$$

   The detection of MN^*_i by a MoS on floor j depends on the vertical coverage of MoSs in floor j.

Under ideal channel conditions, when floor attenuation factor is neglected, RSS measurements only depend on the distance between MNs and MoSs. This implies that, when these two conditions are satisfied, MBFD with $m = 1$ always detects the floor of a MN correctly irrespective of its two dimensional coordinates on the floor.

4.4.2 False Negatives

When a MN is estimated to be on floor j but not on the correct floor i, it is counted as a false negative for floor i, and a false positive for floor j. We call $A_{fn,i}$ – the fraction of areas in floor i declared to be in floor j, $j \neq i$ – as the percentage of false negatives for floor i. In other words, $A_{fn,i}$ denotes the fraction of area in which MNs cannot be located on the correct floor. For good FD performance, $A_{fn,i}$ must be small for all floors i in a given building.

When both constraints mentioned in special case explained in Section 4.4.1.1 are satisfied and channel conditions are ideal, then, a MN in floor i is estimated on the correct floor, and adjacent monitoring ranges (d_{jj} and d_{ji}) cannot influence $A_{fn,i}$.

4.4.3 Effect of Antenna Type

In MBFD, RSS measurements regarding the same MN are collected from multiple floors and the sum of them are compared to each other for the decision. If MoSs do not have adjacent floor coverages, then there will exist RSS measurements from only the floor the MN is residing in. In this case there is no need for FD. However, our objective is to utilize the MoSs across floors in order to increase the overall efficiency of the localization system. Therefore, adjacent (vertical) coverages gain significance in FD performance.

As FD algorithm depends on the adjacent coverages of MoSs, its performance gets affected by the antenna type. The difference between the monitoring range on the vertical direction for isotropic and dipole antennas is expected to cause a difference in floor determination performance for these two types of antennas. With isotropic antennas, although the adjacent monitoring range is smaller than on current floor, the center point of circular coverage is the same on 2 dimensional coordinates. However, with dipole antennas, donut shaped coverage causes an uncovered circular area in the middle of the covered area on the adjacent floor. With both antenna assumptions, an equal size coverage exists on the current floor.

For MBFD to estimate the floor of a MN correctly with $m = 1$, both constraints given in the previous sections must be satisfied. Antenna type related details of 2nd constraint

must be considered carefully depending on the placement of MoSs on adjacent floors.

4.4.4 Evaluation of MBFD Algorithm

According to our simulation results in previous sections, MSAL with 3 floor repetition scheme and 2 floor repetition scheme can provide better localization efficiency when compared to efficiency with SAL. We showed that this is possible without harming the accuracy of localization system. In this section, we provide a detailed performance analysis on FD performance of a 2 floor repetition localization scenario (2FRS) under various channel conditions and with different antenna types. Later in Section A.3.1, the performance of multi-floor subarea localization and multi-floor centroid algorithm with 2FRS are shown when integrated with MBFD algorithm. We have held the exact same evaluation with a 3 floor repetition scenario; however, we chose to present 2FRS evaluation in detail for the sake of simplicity. We provide the evaluation results on 3FRS in Appendix A.3.

While evaluating the floor determination performance, 2FRS $< (5,4), \mathbf{G}_p, A >$ that is introduced in Section 4.3, and shown in Figure 4.4(a) is used. With this placement scheme, for no false negatives, the first constraint in special case is satisfied if $d_{ii} \geq 0.5L$ with both isotropic and dipole antennas. For second constraint to be satisfied, detailed analysis for different antenna types is given as follows:

- Assuming isotropic antennas, the distance MN_i^* to its closest MoS, MoS_i^* on that floor must be upper bounded by the floor height, h, i.e.,

$$min_k \left(dist[MN_i^*, MoS_{ik}]\right) < h$$

  so that MN_i^* is closer to a MoS on the same floor and not one on a different floor.
  This follows from the fact that the minimum distance between a MoS and a MN on the adjacent floor equals to the floor height when they both are located on the same 2D coordinates on adjacent floors. Due to spherical coverage of isotropic antennas, this MoS can monitor the MN. Therefore, the minimum distance, $min_k \left(dist[MN_i^*, MoS_{ik}]\right) = h$. If minimum distance is smaller than floor height, MN can be estimated on the correct floor.

- Assuming dipole antennas, the distance MN_i^* to its closest MoS, MoS_i^* on that floor must be upper bounded as,

$$min_k\left(dist[MN_i^*, MoS_{ik}]\right) < min_k\left(dist[MN_i^*, MoS_{jk}]\right),\ j \neq i$$
$$h^2 > d_s^2 - \sqrt{2} d_s r_{1j}$$

where $d_s = min(dist[MoS_i^*, Mos_j^*])$ in 2D coordinates, so that MN_i^* is closer to a MoS on the same floor and not one on a different floor. This can be explained as follows: According to the placement of MoSs on adjacent floors, we know that the closest distance between two MoSs on the 2-dimensional coordinates is $d_s = dist_{2D}[MoS_i^*, MoS_j^*]$.

$$min_k\left(dist[MN_i^*, MoS_{ik}]\right) < min_k\left(dist[MN_i^*, MoS_{jk}]\right),\ j \neq i$$
$$\sqrt{\frac{r_{1j}^2}{2} + \left(d_s - \frac{r_{1j}}{\sqrt{2}}\right)^2} < \sqrt{h^2 + r_{1j}^2}$$
$$h^2 > d_s^2 - \sqrt{2} d_s r_{1j}$$

Right and left side of this inequality is shown in Figure 4.64 assuming MoSs are placed on a projected grid, and different floor heights. For simplicity, we denote the right side of the inequality with T in the figure. We observe that the condition for zero false negatives is only satisfied when $h \geq 0.35L$ and $d_{jj} < 0.86$. Also, as h is smaller, the monitoring range has to be smaller too.

We show the percentage of false negatives with isotropic antennas when $h = 0.5L$ with 2FRS scenario in Figure 4.65. In this case, it is obvious that when $d_{ii} \geq L/2$, the first constraint is satisfied. When $m = 1$ at $h = 0.5L$, second constraint is also satisfied, and false negatives percentage is 0. As m increases, the range that can provide zero false negatives narrows down because multiple RSS values from each floor are combined together and compared against each other for FD. In 2FRS scenario, for every MN, m RSS measurements are collected from 2 different floors and the sum of these values are compared to each other to find the correct floor among the two floors. From Figure 4.64, we can observe that with dipole antennas, zero false negatives can be achieved when $d < 0.86$ and $h = 0.35L$, assuming d satisfies the first condition. Therefore, in Figure 4.66, we show the A_{fn} on two floors using dipole antennas when $h = 0.35L$. The constraint on d_{jj} is also observed in this figure. We

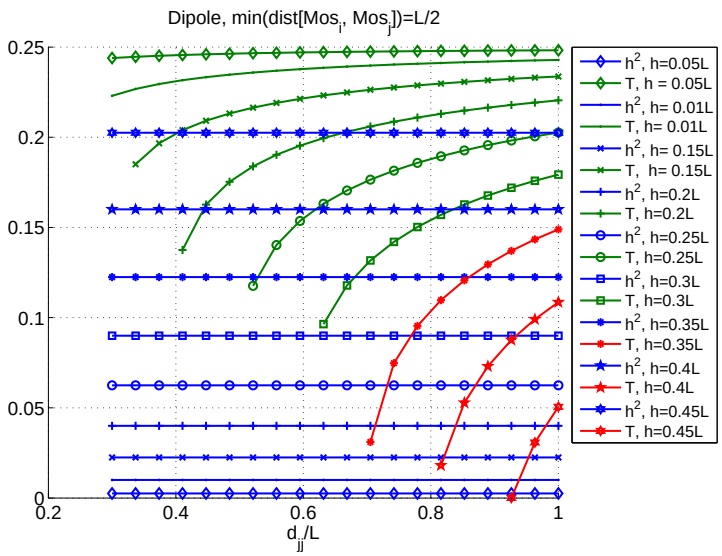

Figure 4.64: Second constraint for dipole antennas

remind that for adjacent monitoring ranges to be non-zero with dipole antennas, $d_{ii} > 2h$ must be satisfied.

We determine the smallest possible fraction of false negatives A^*_{fn} over the possible monitoring ranges in each floor. Figures 4.67 and 4.68 show the *minimum* average percentage of false negatives, A^*_{fn} that can be achieved for a given decision depth m, floor height, h, and the standard deviation of the shadowing, σ with 2FRS scenario, where:

$$A^*_{fn} = min_{d_{ii}, i=1,2} \left(\frac{1}{2} \sum_{i=1}^{F} A_{fn,i} \right). \tag{4.19}$$

In these figures, we ensure that the 1st constraint is satisfied, i.e., under ideal conditions, full coverage on every floor is provided by adjusting the monitoring ranges. In general, we observe increasing percentage of false negatives for increasing floor height and shadowing effect. The effect of floor height arises from the adjacent monitoring range depending on floor height as discussed before. Shadowing effect disrupts the adjacent monitoring range, therefore, false negative percentage gets higher. With dipole antennas, we observe $A^*_{fn} = 0$ when $h \leq 0.35L$ as discussed previously. Monitoring ranges are allowed to be $d_{ii} > 2h$ so

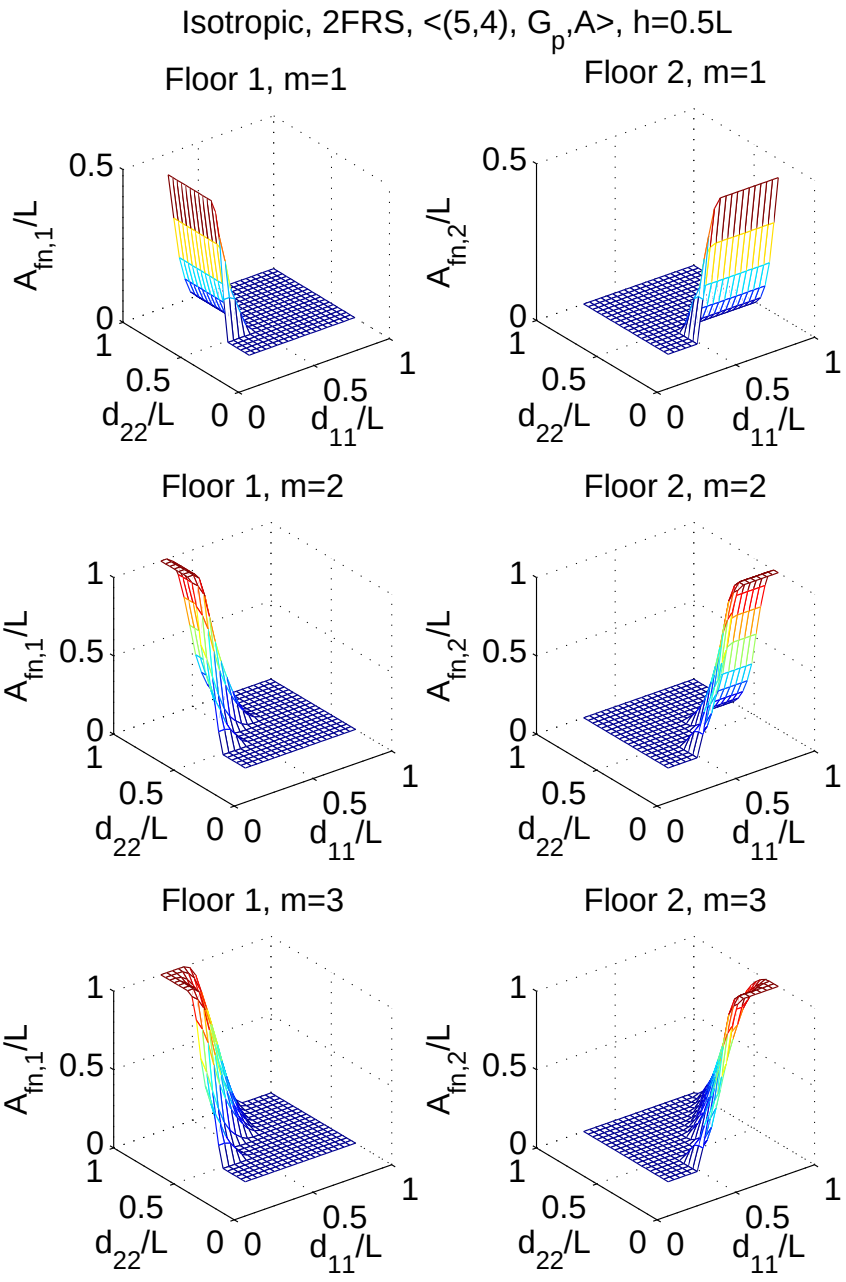

Figure 4.65: Percentage of false negatives with $< (5,4), \mathbf{G}_p, A, iso >$, $\sigma = 0$, $h = 0.5L$

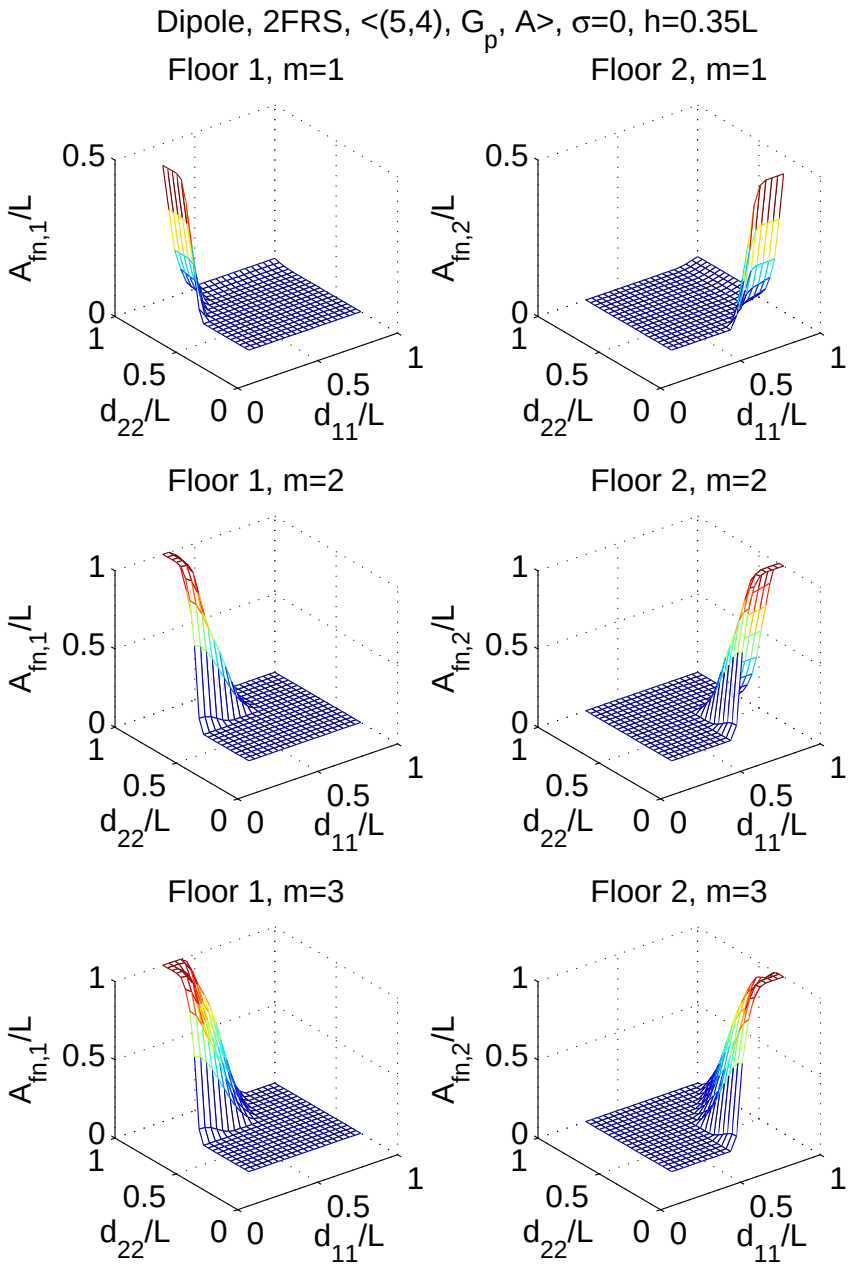

Figure 4.66: percentage of false negatives with $< (5,4), \mathbf{G}_p, A, iso >$, $\sigma = 0$, $h = 0.25L$

that MSAL can be utilized. When this requirement is removed, we do not see a difference in FD performance with isotropic antennas. But with dipole antennas, we see $A^*_{fn} = 0$ when $h \geq 0.3L$. The reason is that at $0.3L$, the first constraint is satisfied, but the monitoring range $d < 2h$, therefore, adjacent monitoring ranges do not exist, then, in floor determination each MN is detected by MoSs on the floor they reside in. In this case, multiple RSS readings cannot be collected from multiple floors.

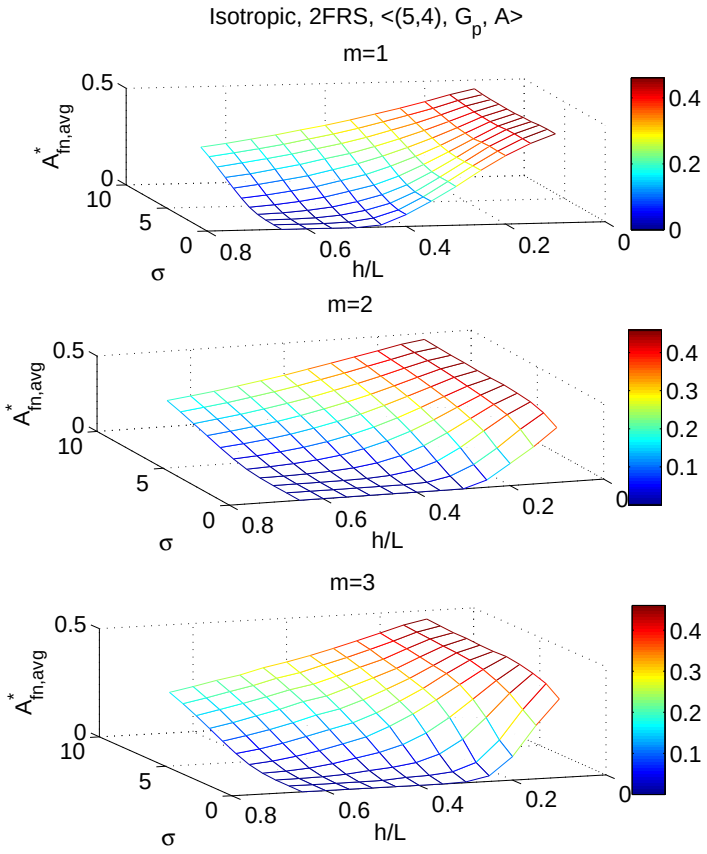

Figure 4.67: Minimum false negative percentage with $< (5,4), \mathbf{G}_p, A, iso >$

To get a more detailed look, Figure 4.69 shows minimum false negatives percentage with respect to floor height under ideal conditions. According to analysis, percentage of false negatives is 0 at decision depth $m = 1$ when $h \geq 0.5L$ and $\sigma = 0$. As also shown with the conditions, at $h = 0.5L$, decision depth $m = 1$ can provide zero false negatives under ideal conditions. However, using a larger decision depth can provide a lower A_{fn} value for smaller

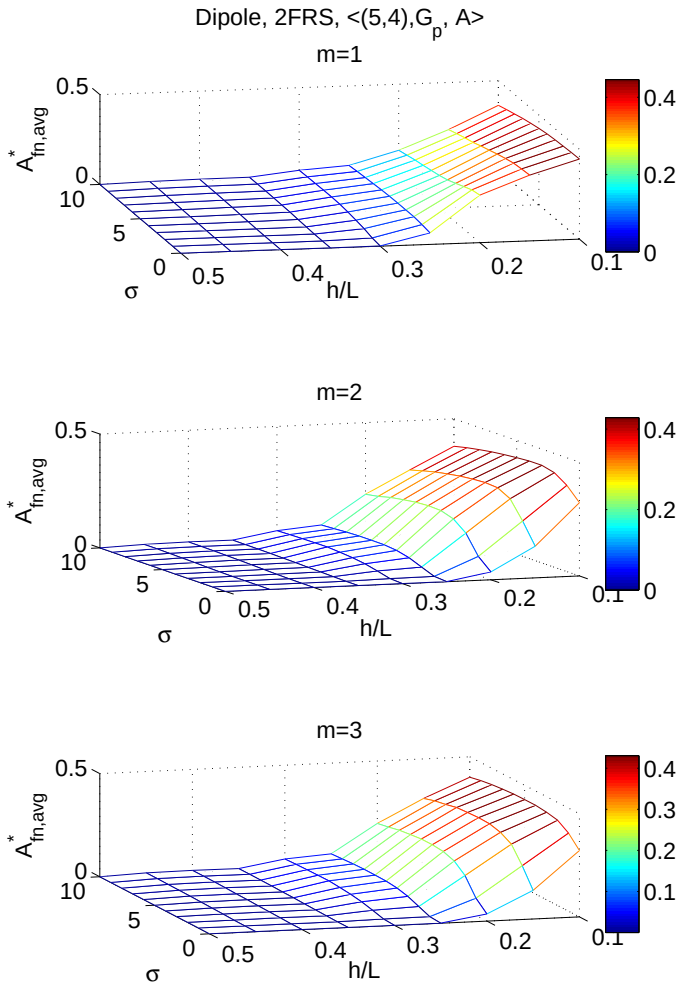

Figure 4.68: Minimum false negative percentage with $< (5,4), \mathbf{G}_p, A, dpl >$

floor height. We also observe that floor height has a sharper effect on false negatives with dipole antennas. This is due to smaller vertical range of dipole antennas.

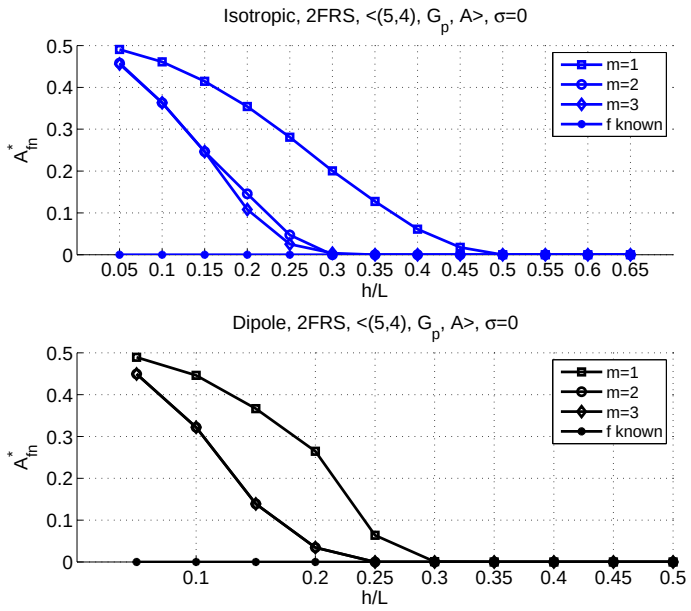

Figure 4.69: Minimum false negative percentage with $< (5,4), \mathbf{G}_p, A, iso/dpl >$, $\sigma = 0$

Figure 4.70 shows the minimum false negatives with respect to σ when $h = 0.25L$. With isotropic antennas, we observe A^*_{fn} with decision depth $m = 1$ is higher than with $m = 2, 3$ when $\sigma < 5$. Under higher shadowing effect, $m = 1$ gives the smallest percentage of false negatives. A^*_{fn} with dipole antennas is much larger than A^*_{fn} with isotropic antennas due to smaller vertical range on dipole antennas when compared to isotropic antennas. As σ is getting larger, we observe A^*_{fn} continues to increase, and smaller A^*_{fn} is given by decision depth $m = 2, 3$ when $\sigma \leq 8$.

4.5 EVALUATION OF MSAL WITH MBFD

Floor determination is independent of the 2D localization algorithm that is used to find where on a given floor the MN is located at. However, it affects the overall performance

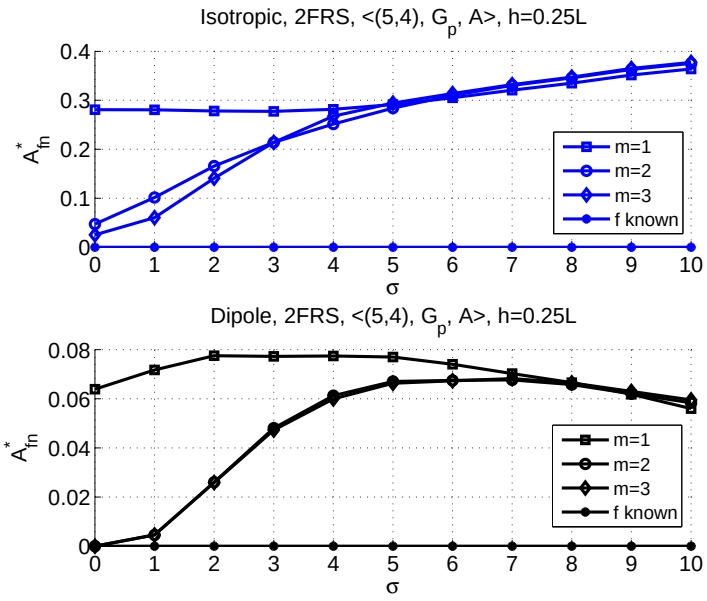

Figure 4.70: Minimum false negative percentage with $< (5,4), \mathbf{G}_p, A, iso/dpl >$, $h = 0.25L$

of localization system in terms of efficiency. Our definition of efficiency comprises of good accuracy, high availability and small number of infrastructure elements per unit area. Because FD changes the availability of location estimates, we want to investigate the effect of different decision depths and antenna types with 2FRS, $< (5,4), \mathbf{G}_p, A >$. This scenario has been used in determining the floor determination performance of majority based floor determination algorithm in the previous section.

4.5.1 Availability with MBFD

In considering the efficiency, we would like to consider only nodes that can be localized, and that too on the correct floor. Availability definition in Section 4.1.3 takes the invalid and uncovered areas into consideration. In this section, we also must consider $A_{fn,i}$ which corresponds to the fraction of area in which MNs cannot be located on the correct floor. Towards this, we make the following definitions and computation.

Let C_i' be the area on floor i that can be sensed by at least one MoS. Then, $C_i'(1 - A_{fn,i})$

is the area on floor i where a MN is determined to be on the correct floor. When $\sigma = 0$, MNs are always located in valid sub-areas whose boundaries are determined from the unit disk model. When $\sigma > 0$, an area $C_{invalid,i}$ of the localizable area $C'_i(1 - A_{fn,i})$ may belong to invalid subareas. As discussed in Section 4.1.3, these arise because radio propagation vagaries will create intersections of monitoring areas of MoSs, that theoretically do not have intersecting ranges. MSAL cannot be directly used for localizing nodes in these invalid subareas. Nodes in such invalid subareas can be localized for instance, by using MCA. Let $C_i = C'_i(1-A_{fn,i}) - C_{invalid,i}$. Then, the fraction of the area in which a node can be localized by MSAL on a floor i is

$$R_{MSAL,i} = \frac{C_i}{|A|}. \tag{4.20}$$

Since MCA can localize nodes in invalid subareas, the availability with MSAL+MCA and MCA are the same and calculated as,

$$R_{MSAL+MCA,i} = R_{MCA,i} = \frac{C'_i(1 - A_{fn,i})}{|A|}. \tag{4.21}$$

Under ideal conditions and when floor information is known, availability is not affected by shadowing or the performance of FD algorithm, but affected only by MoS coverage. MoS coverage depends on floor height as adjacent monitoring range is a function of vertical coverage. When floor information is not given, we use the FD algorithm with different decision depths, m. In the following we show the average availability in 2FRS scenario with different antenna type assumptions when floor information is available and when floor determination is used with $m = 1, 2, 3$. We note that under ideal conditions, $C_{invalid,i} = 0$ and $R_{MSAL,i} = R_{MSAL+MCA,i} = R_{MCA,i} = R_i$. Average availability is given as $R_{avg} = \frac{1}{2}\sum_{i=1}^{2} R_i$.

Average availability on two floors are shown in Figures 4.71 and 4.72 using isotropic and dipole antennas. We assume $h = 0.25L$. When $h < L/2$, floor determination algorithm causes the availability to decrease below 100% even under ideal channel conditions and monitoring range on current floor is large enough to cover the whole area on both of the floors. This was shown before in Figure 4.70. With dipole antennas in order for floor determination can be realized between multiple floors, $d > 2h$. Therefore, we can neglect $d < 0.5L$ when $h = 0.25L$, and focus on where $d \geq 0.5L$. Figures A6 and A7 show the availability that can be achieved with isotropic and dipole antennas on two floors.

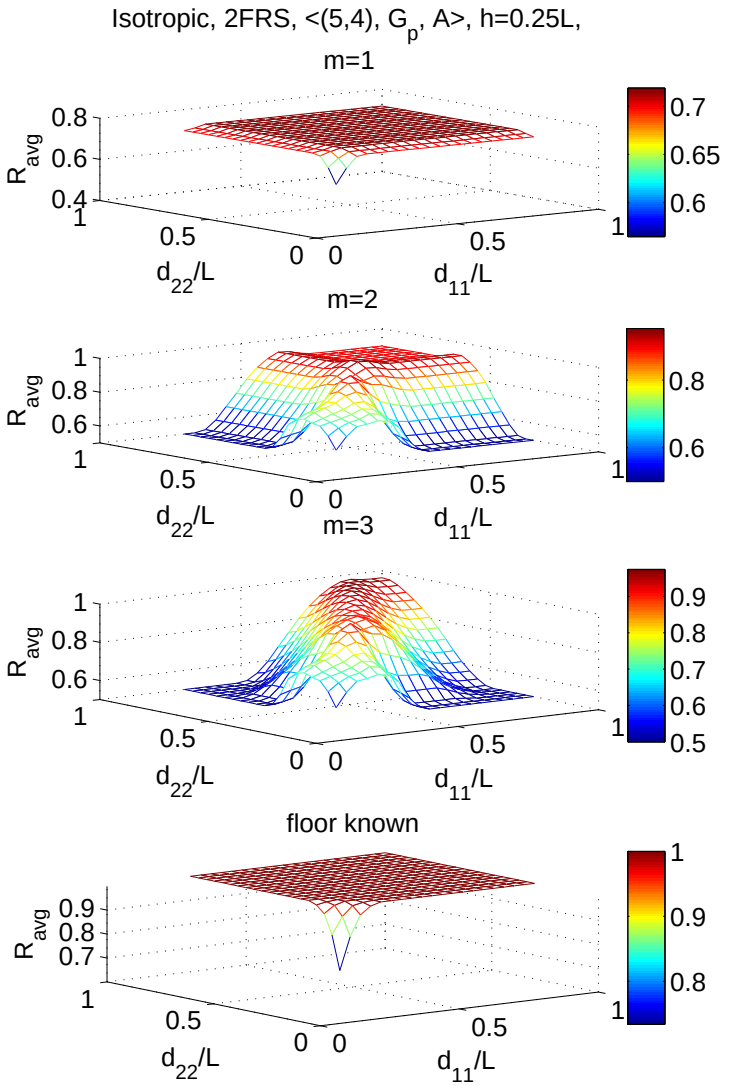

Figure 4.71: 2FRS Average availability with different decision depths with isotropic antennas, $\sigma = 0$, $h = 0.25L$

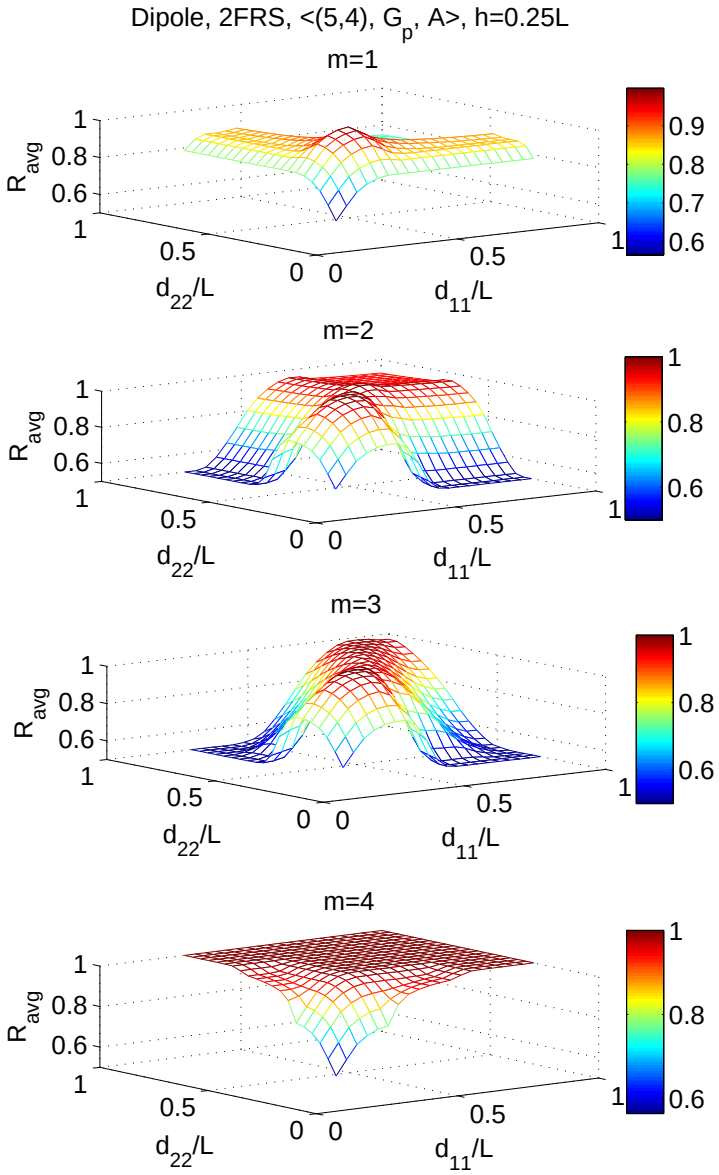

Figure 4.72: 2FRS Average availability with different decision depths with dipole antennas, $\sigma = 0$, $h = 0.25L$

4.5.2 Best Accuracy Conditions

Depending on floor determination algorithm's performance with different decision depths, the availability of location estimation increases or decreases. This reflects as higher or lower minimum average localization error.

Figure 4.73 shows the accuracy with different floor heights under ideal channel conditions with 2FRS. With both MSAL and MCA, $m = 3$ is providing the best accuracy levels, while $m = 1$ has the worst accuracy levels. This is due to availability degradation with larger decision depth. Decision depth $m = 1$ provides the closest accuracy levels to the levels with known floor information. When $h \geq 0.55L$, we observe that CA-iso is preferred instead of MCA-iso.

With dipole antenna assumption, best achievable accuracy approaches to the same value for different decision depths at $h \geq 0.35L$ as shown in Figure 4.73. The reason can be explained as follows: When floor height is larger than a threshold, in order to use adjacent monitoring, monitoring ranges must be increased; however, increasing monitoring ranges may not be the best choice in accurate location estimation of MNs on the current floor. Therefore, when floor height is larger than this threshold, instead of using MSAL, SAL is used in location estimation. In this case, there is no need to floor determination algorithm and the percentage of false negatives is zero. Therefore, the availability is the same as in the case when floor information is known.

Effect of log normal shadowing with $\sigma = 5$ on localization accuracy of 2FRS is given in Figure 4.74 when different decision depths are used in FD. Our first observation is that the difference in accuracy caused by different decision depths is disappearing under shadowing with $\sigma = 5$ when MSAL+MCA or MCA is used. When MSAL is used in localization, $m = 1$ provides the closest accuracy to the accuracy with known floor information. Decision depths $m = 2$ and $m = 3$ causes worse availability therefore lower localization errors. Among the three localization algorithms MSAL performs best in terms of accuracy although availability with MSAL+MCA is higher. Error with $m = 3$ is the smallest due to higher percentage in false negatives (lower availability) at this decision depth.

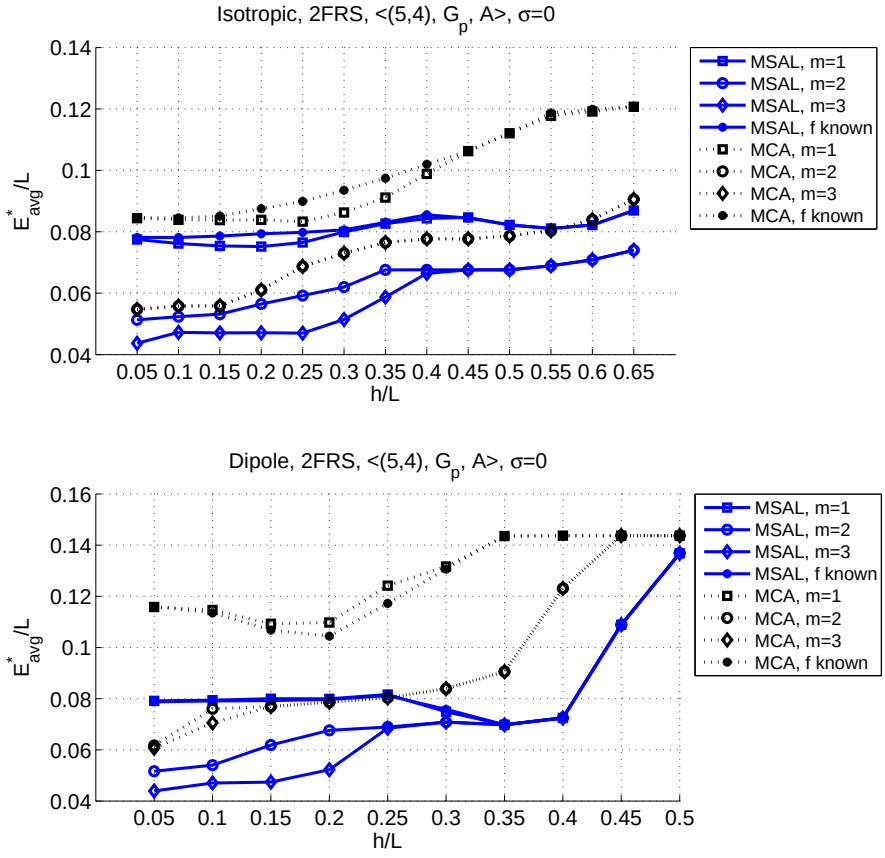

Figure 4.73: Minimum localization error with respect to h when $\sigma = 0$, 2FRS

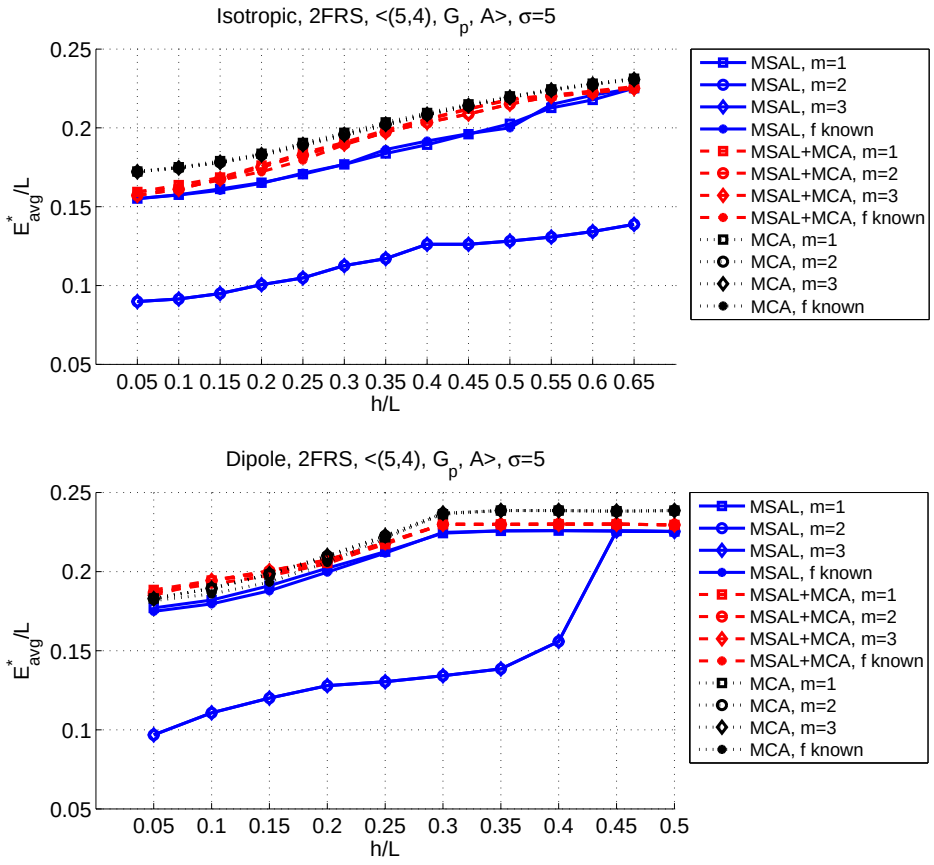

Figure 4.74: Minimum localization error with respect to h when $\sigma = 5$, 2FRS

4.5.3 Best Efficiency Conditions vs. Floor Height

Efficiency with known floor information constitutes the upper bound of efficiency that can be achieved with any floor determination algorithm under ideal conditions.

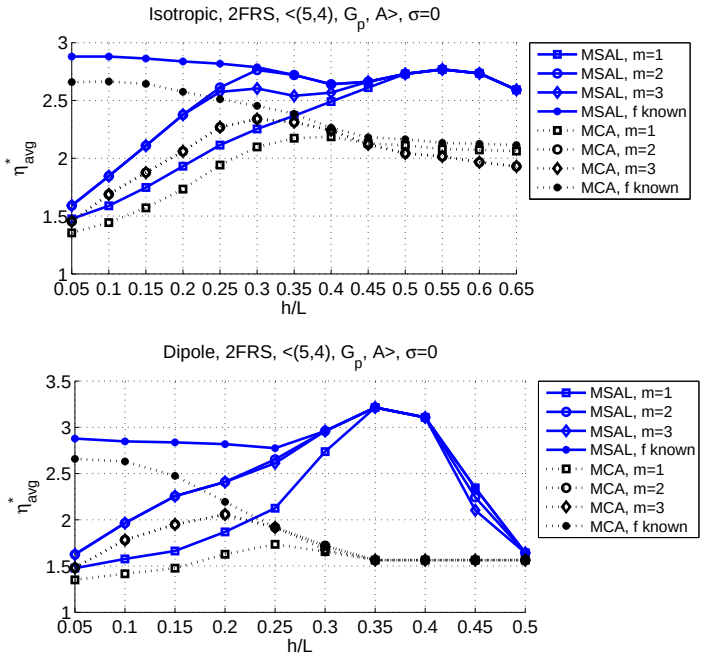

Figure 4.75: Maximum efficiency with respect to h when $\sigma = 0$, 2FRS

Figure 4.75 shows the maximum efficiency that can be achieved under ideal channel conditions with 2FRS. Maximum efficiency under ideal conditions is achieved by MSAL. As floor height increases, the maximum efficiency that can be achieved with FD approaches to maximum possible value that can be achieved assuming known floor information, therefore no false negatives. A_{fn} decreases as the floor height increases, this reflects as better availability and better efficiency. When $h > 0.5L$, availability stays at 100%; however accuracy gets affected by large floor height. Therefore, worse efficiency levels are seen. Although it is not possible to change the floor height, it looks like there is an optimum value for h that is maximizing the efficiency. For isotropic coverage and dipole coverage, $h = 0.55L$ and $h = 0.35L$ are the best values. Accuracy at maximum efficiency levels with different decision depths are similar to each other. As floor height increases, SAL is preferred to MSAL on

one of the floors; as h continues to increase SAL is preferred on both floors. If d_{ii} is set so that only MSAL is allowed, then we would see with this is only possible when $h \leq 0.3L$ and $h \leq 0.1L$ with with isotropic and dipole antennas, respectively.

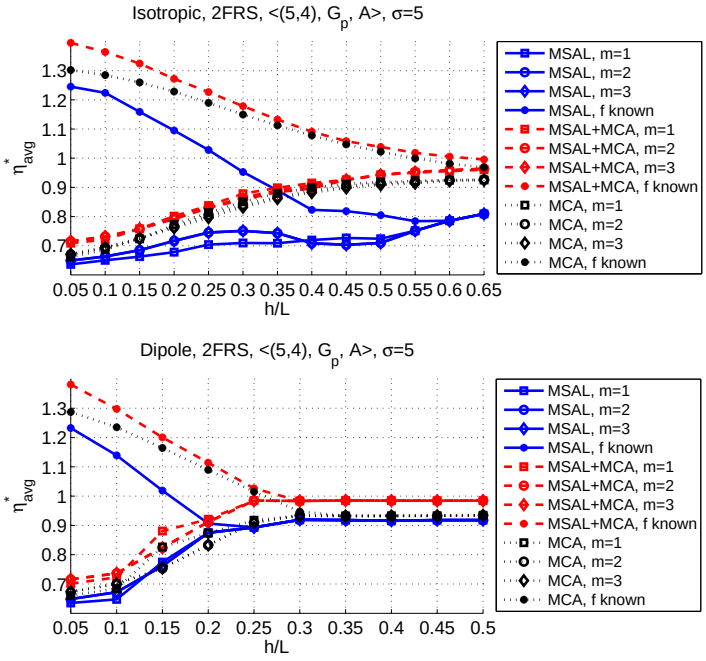

Figure 4.76: Maximum efficiency with respect to h when $\sigma = 5$, 2FRS

Figure 4.76 shows the maximum efficiency that can be achieved under shadowing with $\sigma = 5$ with 2FRS. Maximum efficiency is achieved by MSAL+MCA with known floor information. As floor height increases, maximum efficiency with floor determination approaches to maximum efficiency level. Efficiency levels with different decision depths are similar to each other under this shadowing level.

4.5.4 Best Efficiency Conditions vs. Shadowing Effect

Finally, we show the effect of shadowing effect on maximum efficiency with 2FRS using floor determination algorithm with various decision depth values.

Figures 4.77, 4.78 and 4.79 show the maximum efficiency, average localization error and availability possible at maximum efficiency conditions with respect to σ when $h = 0.25L$

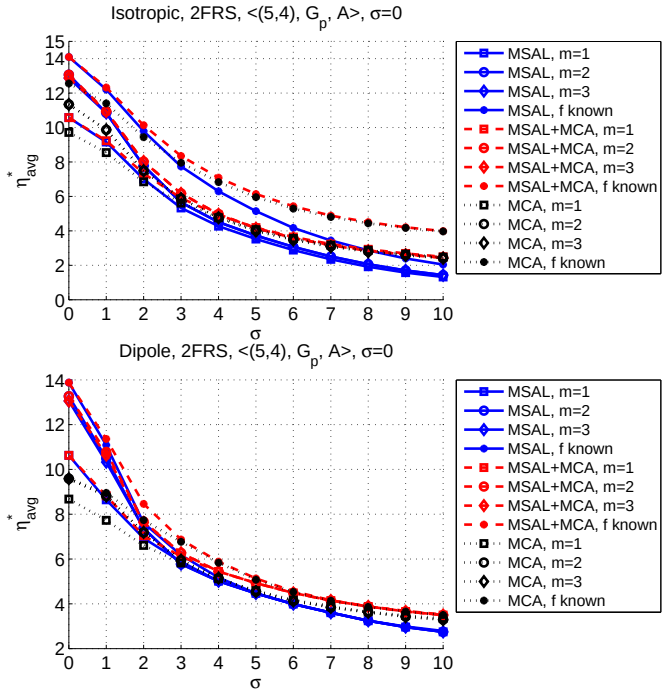

Figure 4.77: Maximum efficiency with respect to σ when $h = 0.25L$, 2FRS

with 2FRS. We observe the decision depth effect on maximum efficiency is minimal especially under high shadowing effect. With dipole antennas, when $\sigma \geq 6$, we see similar efficiency with floor information is known or estimated. The reason is that in these cases, SAL is preferred to MSAL.

Figure 4.78 shows the average localization error at maximum efficiency conditions. At small σ values MSAL and MSAL+MCA perform similarly and outperform MCA, whereas at high σ values MSAL provides the smallest localization error due to lower availability. The observation is similar for both isotropic and dipole antenna assumptions.

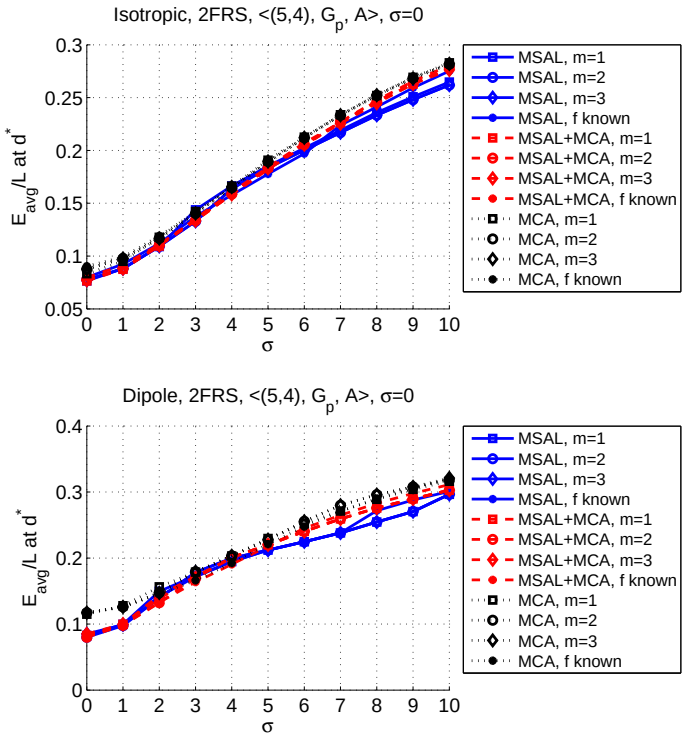

Figure 4.78: Average localization error at maximum efficiency conditions with respect to σ when $h = 0.25L$, 2FRS

Figure 4.79 shows the availability at maximum efficiency conditions. This figure also explains the degradation in maximum efficiency with isotropic antenna assumption. We observe a decrease in availability due to false negatives caused by floor determination algorithm.

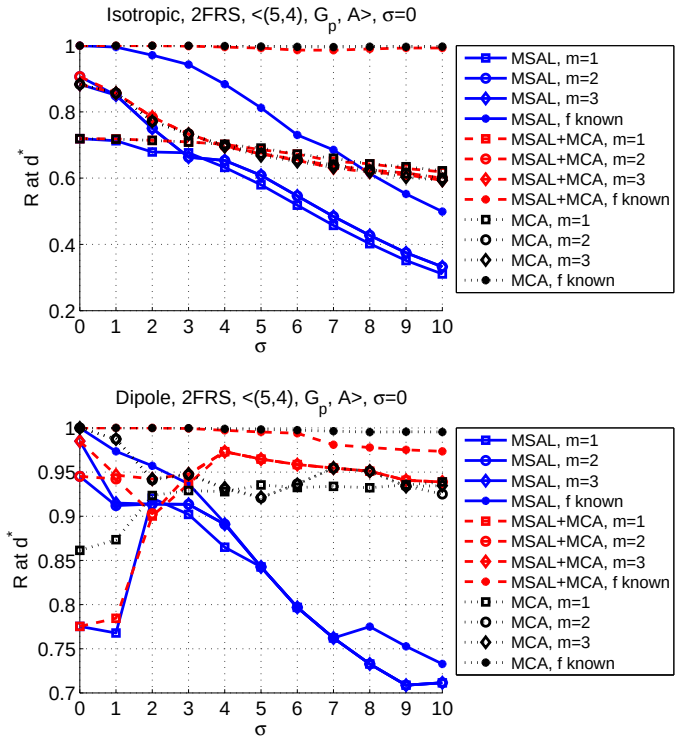

Figure 4.79: Availability at maximum efficiency conditions with respect to σ when $h = 0.25L$, 2FRS

4.5.5 Effect of Floor Attenuation Factor

Floor attenuation factor (FAF) affects the way the floor determination and localization algorithms perceive the floor height. When FAF is utilized in signal propagation model, floor height is perceived as if it is larger than it actually is. Therefore, we expect the minimum localization error and maximum efficiency values using different decision depths will merge to the same values at smaller floor height values than the values shown in previous section, where we did not consider FAF.

4.5.5.1 Best Accuracy Conditions Figure 4.80 shows the minimum localization error with 2FRS under ideal channel conditions. When we compare this figure to Figure 4.73,

we observe that the localization errors with different decision depths are approaching to the same value at a smaller floor height, i.e., $h \geq 0.35L$ and $h \geq 0.2L$ with isotropic and dipole antenna assumptions, respectively. At these floor heights, we see similar accuracy values with isotropic and dipole antennas. The reason is that SAL is preferred to MSAL, and on current floor isotropic and dipole antennas has the same coverage. For h smaller than these values, localization error with FAF is higher than without FAF. The reason is the narrower adjacent monitoring ranges because of larger path loss due to floor attenuation.

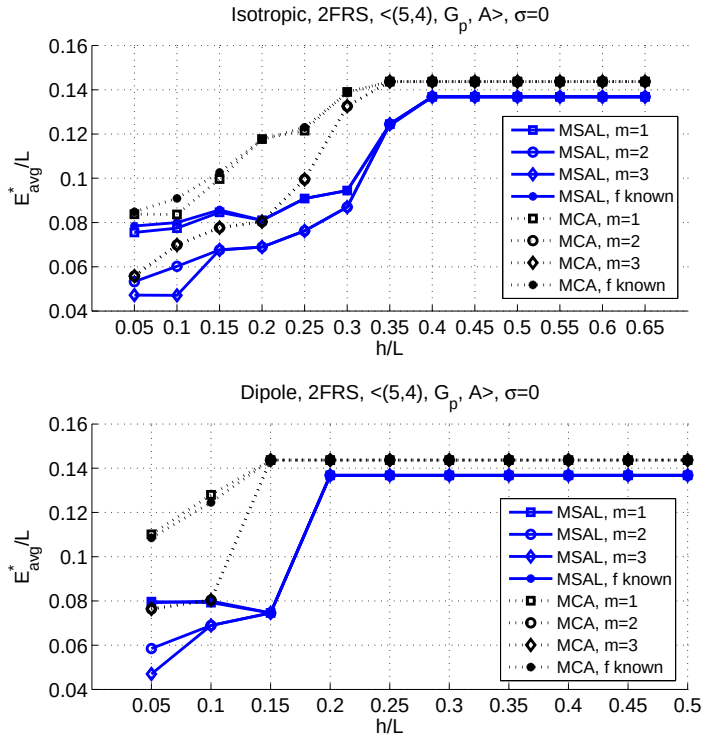

Figure 4.80: Minimum localization error with respect to h when $\sigma = 0$, 2FRS

Figure 4.81 demonstrates the same results from the false negatives point of view with 2FRS under best accuracy conditions. False negatives percentage approaches to zero as floor height increases.

Similar conclusions can be drawn from Figure 4.82 showing the minimum localization error under shadowing with $\sigma = 5$. As shown in Figure 4.83, false negatives percentage is higher under shadowing effects, and smallest with $m = 1$.

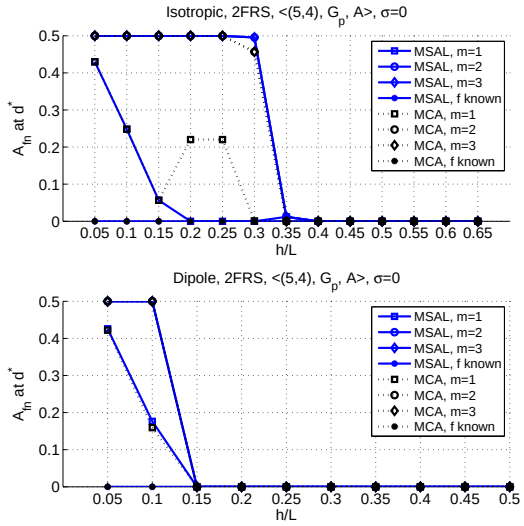

Figure 4.81: False negative percentage with respect to h when $\sigma = 0$, 2FRS

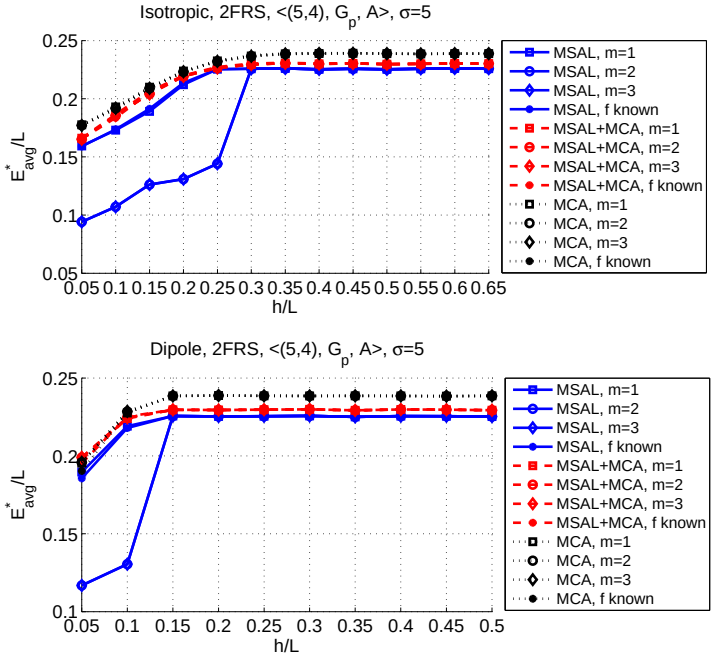

Figure 4.82: Minimum localization error with respect to h when $\sigma = 5$, 2FRS

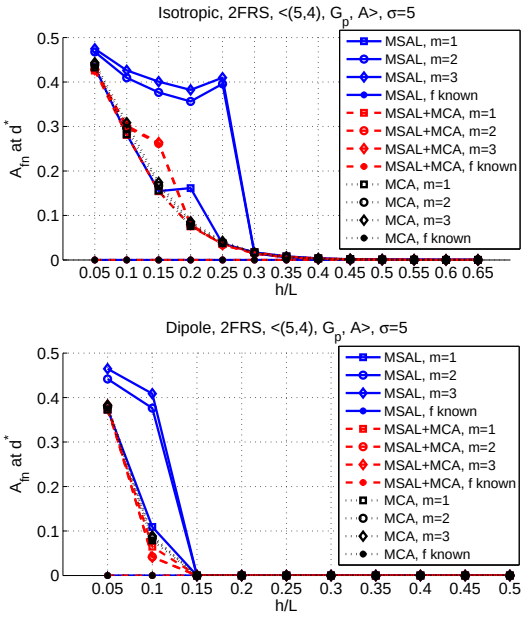

Figure 4.83: False negative percentage with respect to h when $\sigma = 5$, 2FRS

4.5.5.2 Best Efficiency Conditions vs. Floor Height The best efficiency that can be achieved with 2FRS under ideal channel conditions with FAF is shown in Figure 4.84. When compared to Figure 4.75, we do not observe a significant decrease in efficiency values with isotropic antennas at small floor height. However, we do observe that at larger floor height, the efficiency with MSAL approaches to efficiency with MCA. This implies that at these floor height values, adjacent monitoring ranges approaches to zero, and SAL is used for localization. Since in SAL only current floor MoSs are used, the number of MoSs used in localization decreases, therefore accuracy gets worse as so efficiency. This is the reason why the efficiency approaches to the same values when $h \geq 0.4L$ and $h \geq 0.2L$ with isotropic and dipole antenna, respectively.

Best efficiency under shadowing with $\sigma = 5$ is shown in Figure 4.85. When compared to the case without FAF shown in Figure 4.76, we observe that efficiency with and without floor determination approaches to same values at a smaller floor height as expected. Also, when floor heights is larger than these values, we observe similar results as SAL is used in

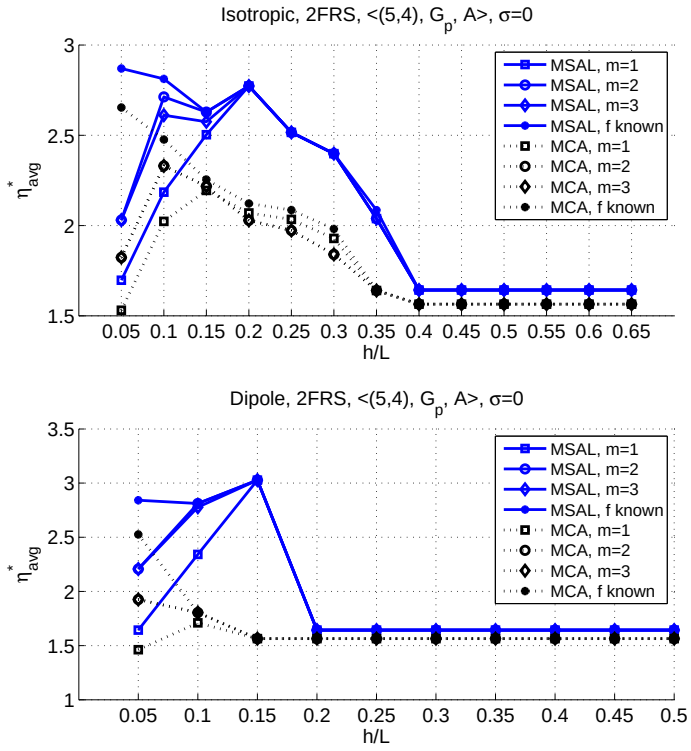

Figure 4.84: Maximum efficiency with respect to h when $\sigma = 0$, 2FRS

localization instead of MSAL.

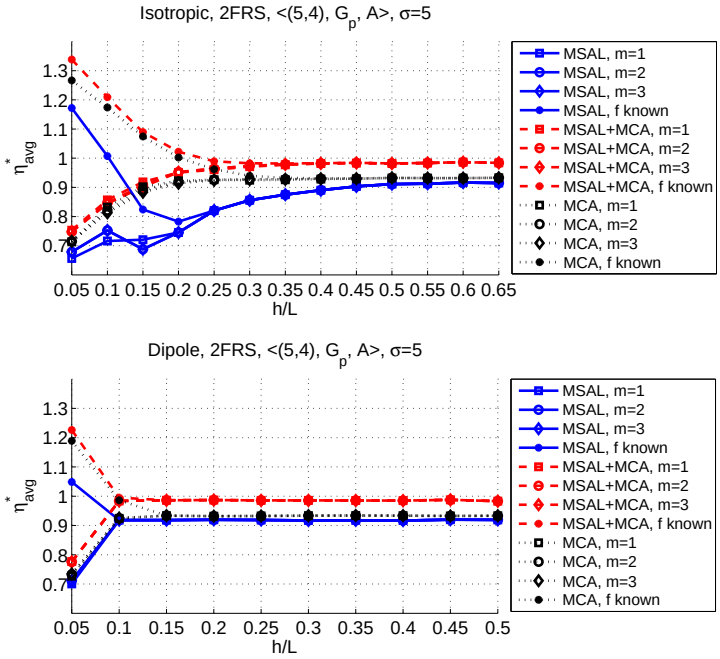

Figure 4.85: Maximum efficiency with respect to h when $\sigma = 5$, 2FRS

Figures 4.86 and 4.87 show the availability during best efficiency conditions under ideal conditions and shadowing with $\sigma = 5$. As also observed from the figures, we allow the coverage to be less than 100%. This is why at large floor height values, availability is not 100%. In addition, we observe higher availability at larger floor height, SAL is started to be used rather than MSAL.

4.5.5.3 Best Efficiency Conditions vs. Shadowing Effect Lastly, we want to present the best efficiency conditions under shadowing with varying σ assuming $h = 0.25L$. Figures 4.88 and 4.89 show the best efficiency levels and localization accuracy available at these levels with and without floor determination. When compared to the case without FAF shown in Figure 4.77, we observe a degradation in efficiency especially at lower σ values. At higher values, the difference vanishes. In both cases, decision depth effect is insignificant, and MSAL+MCA outperforms other algorithms.

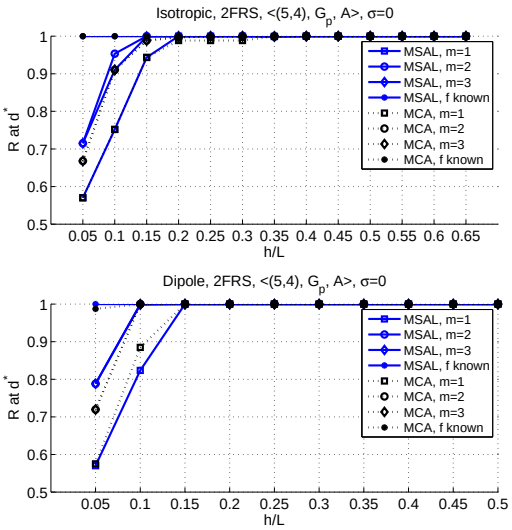

Figure 4.86: Availability with respect to h when $\sigma = 0$, 2FRS

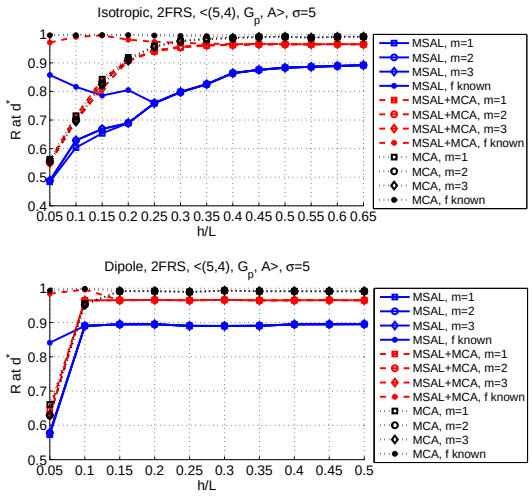

Figure 4.87: Availability with respect to h when $\sigma = 5$, 2FRS

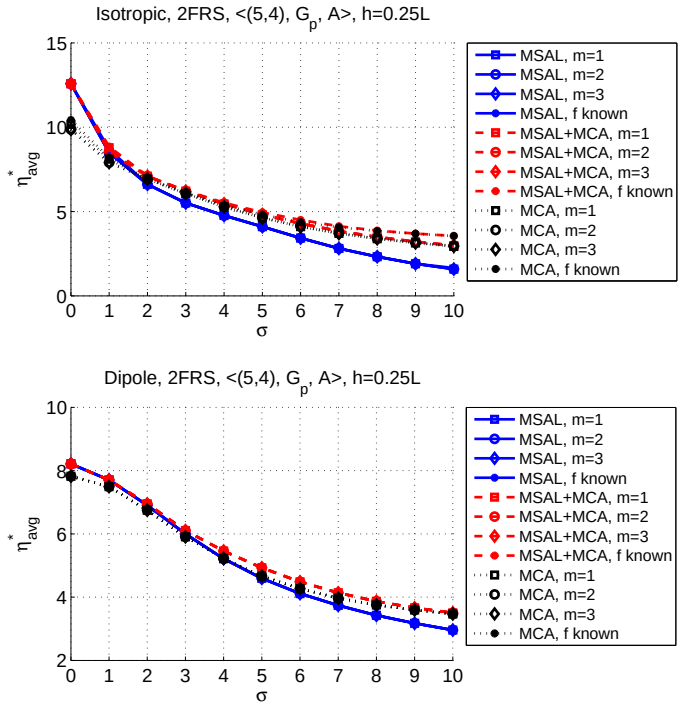

Figure 4.88: Maximum efficiency with respect to σ when $h = 0.25L$, 2FRS

From Figure 4.89, we observe MSAL+MCA and MSAL can provide best accuracy with dipole antennas for the whole range of σ values. And, with isotropic antennas, MSAL+MCA and MSAL perform similarly at low σ values, and as σ increases localization error of MSAL+MCA approaches to error with MCA. The reason is obviously the availability and false negatives percentage effects. False negative percentages with two types of antennas are shown in Figure 4.90.

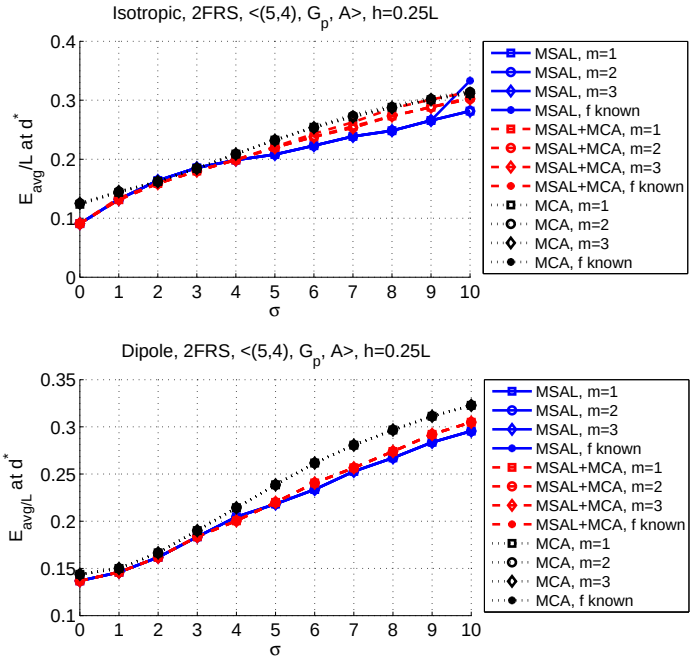

Figure 4.89: Localization error with respect to σ when $h = 0.25L$, 2FRS

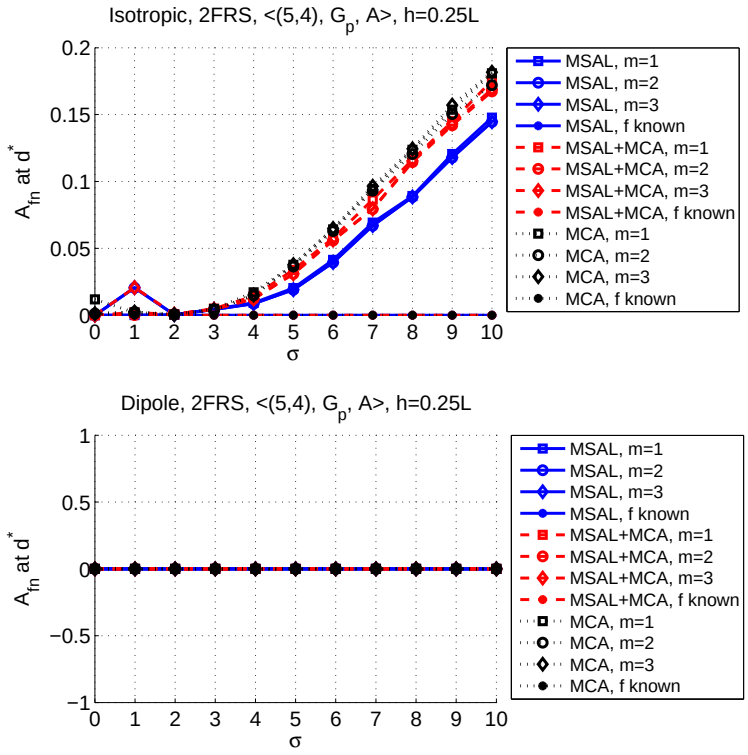

Figure 4.90: False negative percentage with respect to σ when $h = 0.25L$, 2FRS

4.6 CONCLUSION

We have introduced a proximity based multi-floor localization approach that increases deployment efficiency by utilizing MoSs in multiple floors. We investigated the effect of antenna model, floor height, floor repetition scheme and the variance of the shadow fading on the efficiency of the system. Our simulation results indicate that using MoSs across floors can increase the deployment efficiency by up to 50%. Our multi-floor localization system provides within-floor localization accuracy comparable to that on a single floor with fewer MoSs. Under ideal conditions, when best monitoring ranges are determined, isotropic and dipole antenna coverages leads to similar localization error performances using MSAL with both 2 and 3 floor repetition schemes. The best ratio of monitoring ranges of MoSs on adjacent floors is between 0.5 and 1.5 depending on floor height and shadow fading. However, as RSS shadowing fluctuations are included, system designers can select the same monitoring range on all floors without significantly sacrificing accuracy. The performance in real systems may be tuned by changing signal sensing thresholds, using attenuators. Estimating channel parameters through online calibration can benefit in determining the monitoring range close to reality. This is part of our ongoing work. Finally, in real environments, several (sometimes > 10) WiFi access points are often visible at many locations. These access points are placed randomly but may create subareas suitable for localization. We plan to experimentally evaluate the localization performance in these scenarios.

The key takeaways from the tables summarizing our findings in A.1.3 are:

- As shadowing effects gets more powerful, accuracy and efficiency get worse with every localization scenario and every localization algorithm employed.
- Different floor plans with different wall or furniture characteristics will result in different accuracy and efficiency conditions with different best monitoring ranges on different floors. However, it has been observed that there is a minor difference between best accuracy and best efficiency levels on two floor plans that we have simulated with different localization scenarios, especially under shadowing effect.
- FAF changes how localization algorithm perceives the floor height. Therefore, a significant decrease is observed both in efficiency and in accuracy.

- With MSAL+MCA, efficiency is higher than with MSAL only localization. The reason is the availability of location estimates for MNs detected in invalid subareas with MSAL+MCA. A detailed discussion has been provided in the previous sections.
- With random placement of the same total number of MoSs on two floors, we have shown that the accuracy gets worse by 0.06L and 0.03L under ideal conditions and under shadowing with $\sigma = 5$, respectively. This small difference implies that the accuracy will not suffer significantly when MoSs cannot be placed on a projected grid.
- Under ideal conditions, MSAL and MSAL+MCA performs exactly the same in all localization scenarios except the ones including WAF effect. The reason is that the localization system assumes signals arriving to a MoS receiver passes through the same n_{avg} number of walls while traveling. However, in reality, they may be passing through a less or more number of walls. This causes the creation of invalid subareas even under ideal channel conditions shown with $\sigma = 0$. While MSAL+MCA can localize MNs detected in these invalid subareas, MSAL cannot; therefore, differences in performance occurs. signals arriving to a MoS receiver passes through the same n_{avg} number of walls while traveling. However, in reality, they may be passing through a less or more number of walls. This causes the creation of invalid subareas even under ideal channel conditions shown with $\sigma = 0$. While MSAL+MCA can localize MNs detected in these invalid subareas, MSAL cannot; therefore, differences in performance occurs.
- When considering the performance with dipole antennas, we have to keep in mind that for multi-floor localization to be used, $d > 2h$ ($d > 2h_{FAF}$ with scenario including FAF). Therefore, in order to use multi-floor localization, monitoring range must be larger than the monitoring range that is required with isotropic antennas. We can see this in Tables A5 and A13 for 2FRS and in A7 and A15 for 3FRS. Under ideal conditions, multi-floor localization is preferred to single floor localization as can be inferred from optimum monitoring ranges. However, increasing the monitoring range so that the adjacent coverages will exist may result in not optimum results from a single floor point of view. Therefore, we also observe when dipole antennas are utilized, single floor localization is preferred to multi-floor localization. These are observed with WAF and FAF scenarios under shadowing effects. This statement is valid for both 2FRS and 3FRS.

- In terms of localization efficiency, dipole antennas can perform as good as isotropic antennas systems under ideal conditions and open environment with no FAF and WAF effect. However, as shadowing starts to take effect or WAF and FAF are included, a significant degradation in efficiency is observed. Nevertheless, the performance with MSAL+MCA is comparable or better than multi-floor centroid algorithm.
- Although floor determination performance does not depend on localization algorithm, it does depend o placement of Moss, floor height, and floor repetition schemes in multi-floor buildings. On the other hand, it affects the performance of localization system in terms of both accuracy and efficiency, because it changes the availability of location estimates. When floor determination algorithm cannot determine A_{fn} percentage of MNs located on a floor in correct floor, then the efficiency of localization system goes down even if those nodes can get location estimates, since the MNs are determined to be on the wrong floor, there is no value of that location estimate.
- In Tables in Appendix A.1.3, we present simulation results of MSAL, MSAL+MCA and MCA accuracy and efficiency conditions under various localization scenarios, and also false negatives with decision depth equals to 1. We remind that decision depth 1 compares maximum RSS values collected from multiple floors to decide in which floor the MN is residing in. For example with MSAL and MSAL+MCA in 2FRS scenario, at the best localization error conditions, we observe 28% false negatives even under ideal channel conditions, and larger than 30% under shadowing.
- When floor attenuation factor is included in signal propagation model, floor height $h = 0.25L$ is perceived as $h_{FAF} = 0.7L$. This corresponds to a special case where $h > 0.5L$, and leads to zero false negatives percentage under ideal channel conditions. And this can also be observed from simulation results. While shadowing starts to take effect, false negatives starts to appear.
- With 3FRS, we observe higher false negative percentage at best accuracy and efficiency conditions. One important reason is that in 3FRS floor determination works to differentiate between three floors whereas in 2FRS only RSS from 2 floors are compared to each other.
- With dipole antennas, we notice false negatives percentage decreases to the contrary

with isotropic antennas with increasing shadowing level. However, when we look at closely, we observe that under high levels of shadowing, SAL is preferred to MSAL or monitoring range is decreased so that adjacent floor coverage narrows down. This results in MNs to be mostly heard by MoSs on their on floor; in effect, a smaller percentage of false negatives. When SAL is preferred, there is no need to floor determination, false negatives do not exist, localization errors are higher, and efficiency is lower.

5.0 CASE STUDY: USING SAL FOR SPATIO-TEMPORAL SENSING

According to FCC's Spectrum Policy Task force, at any given time and location, most of the licensed spectrum is idle. Spectrum shortage resulting from static spectrum management policies can be overcome using dynamic spectrum access techniques to identify and utilize the underutilized spectrum [78]. Underutilization exists if the spectrum can accommodate secondary transmissions without harming the operation of the primary user of the band [79]. In Cognitive Wireless Networks (CWN), in order to increase the flexibility and utilize the available spectrum efficiently, opportunistic spectrum access (OSA) of licensed spectrum by secondary nodes is allowed. The region of space-time-frequency in which a secondary use is possible is called a *spectrum hole*. To fill in spectrum holes dynamically, wireless systems must determine where the holes exist in space, time, and frequency. Once spectrum opportunities are detected, secondary users must decide whether and how to exploit them (what modulation and transmission power to use, how to share opportunities among secondary users are some of the issues), and reconfigure for opportunistic access to spectrum without harming the primary user. Many studies on methods of spectrum sensing have been proposed for CWN. Energy detection of frequency bands, and event based detection of arrival/departure of signals are examples of two single user algorithms that are mentioned among the others in a survey given in [79]. Cooperation based spectrum sensing was proposed to exploit the diversity gains in wireless medium, soft/hard combining of information, etc [79].

Spatial knowledge of network topology and geometric relations of primary and secondary users can be of significant benefit in detecting spectrum use opportunities. Localization algorithms can be employed to estimate the locations of users in primary and secondary networks, and determine the available spectrum holes. One of the challenges of such local-

ization for identifying spectrum holes is that sophisticated localization techniques may be costly and there may often be no location information readily available. As described previously, Sub-Area Localization (SAL) is a proximity-based localization algorithm which avails of the sub-areas created by the overlapping monitoring coverages of N monitoring stations (MoSs) with known locations. This localization algorithm is explained in detail in Chapter 3. In this chapter, we want to evaluate whether SAL can be utilized for spatio-temporal sensing of primary transmitters with the object of determining spectrum holes.

If the MoSs in SAL are networked together, SAL can provide a global awareness in terms of active primary user topology to secondary users. Without such a system, secondary users will be required to perform local sensing and somehow share this information with each other in order to be aware of transmissions that they are not able to monitor due to their locations and radio propagation vagaries. The sharing process may incur overhead and latency in utilization of spectrum holes. In addition, the information gathered with local sensing greatly depends on the mobility patterns, distribution and the numbers of secondary users. In the case of clustered distribution of secondary users, the sharing process would be redundant, whereas with small number of secondary users, sufficient information regarding primary transmitters may not be collected.

The objective of this case study is to examine the performance of SAL in detecting primary transmitters. We evaluate the efficiency of SAL in determining the spectrum holes under ideal channel conditions. The network may include primary transmitters operating on different frequencies. The localization system can monitor these transmissions, but it however may not be able to modify the "optimum monitoring range" for each different frequency once it is deployed for service. When there is a difference between the frequency of primary users that is being monitored and the frequency for which the localization system is optimized, the estimation accuracy will get affected. Therefore, we also examine the performance of SAL in localizing primary transmitters for detection of spectrum holes when it senses transmissions at frequencies for which it is *not* optimized through simulation results with various path-loss models. Then, we will conclude the chapter with a qualitative discussion on how SAL performance can be perhaps improved under different scenarios such as using secondary network topology information for spectrum hole detection, or the

effect of the location knowledge of primary receivers, etc. We assume that SAL has the RF and antennas necessary for wideband monitoring of spectrum. Our goal is not to provide a solution for spectrum monitoring, but to illustrate the potential of SAL in emerging applications and we employ detection of spectrum holes as an example. Thus, this chapter has its own background section, but the examination of the spectrum hole detection problem is not comprehensive.

5.1 RELATED WORK

In the literature, there exist a number of studies looking at location based spectrum access. These studies argue that the use of network topology information and the geometric relations of the nodes can significantly benefit the operation of CWN. The work in [5] argues that the topology information has direct usage and implications on connectivity and capacity estimates of the network, and should be exploited to optimize network efficiency. The work in [80] focuses on DSA concepts in the context of using spatial statistics to understand the network topology of primary and secondary users. In [81], opportunistic access is allowed only if the experienced interference by the primary licensed users due to the activity of secondary users' does not exceed a predefined threshold. To calculate the transmit power of a secondary user, the positions of primary users must be known so that the interference to primary users can be estimated beforehand and does not exceed the limit. In [82], to find the maximum allowed transmit power of the secondary node while respecting primary constraints, the authors propose a heuristic algorithm that is based on a dynamic threshold for detecting the activity of primary base stations. In this paper, the detection of spectrum holes is based on transmitter detection, considering that the position of primary users are unknown but the boundaries for each cell, where the users are enclosed are approximated and made available to the secondary users.

Radio environment maps (REM) are used to provide an architecture to record, store, access and share relevant information about the environment to assist in functions such as collaborative spectrum sensing [83]. The information stored may be the locations of

transmitters, spectrum usage maps or so on. In [84], the authors proposed an enhanced radio environment map concept that makes use of spatial statistics and probabilistic models to reduce the overhead of basic REMs. Probabilistic models can let the wireless environment be expressed in a very compact form compared to the original data set (perhaps of measured quantities). A topology engine is developed, which is an agent in CWN collecting and processing spatial information about the environment for storage in the REM.

5.2 THE SETTING FOR SECONDARY SPECTRUM USE WITH SAL

We consider a primary network with primary users with an activity pattern that is unknown to the secondary users. Secondary users do not possess spatial topology information of the primary network. We let the spectrum hole be the space, time, and frequency channels that secondary transmitters can transmit. Secondary users can transmit with the highest power possible without causing interference to primary receivers using the same frequency. This is the *primary constraint* that secondary users must satisfy while utilizing the available spectrum. An additional secondary constraint is possible where a secondary transmitter can transmit with the highest power it can transmit without interfering with another secondary user's transmission on same frequency band. While the primary constraint protects the primary network from possible interference from the secondary network, the secondary constraint protects secondary nodes from each other in order utilize the available spectrum fairly.

There are two phases in opportunistic utilization of spectrum holes by secondary nodes. In the first phase, a global spatio-temporal sensing with Sub-Area Localization algorithm is realized to gain knowledge of the topology of primary users. This knowledge along with transmission characteristics will help determine the spectrum holes. In the opportunistic access configuration phase, secondary transmitters receive the knowledge of location estimates for spectrum holes for a given frequency band from the localization system. They use this information to decide whether they can transmit and with which power and on which frequency according to the primary and secondary constraints. We do not consider in detail

the protocols to convey this information or the geographic identification of locations and positions (e.g., whether latitudes and longitudes are used or local coordinates). Similarly, other details (e.g., the need and use of common channels, perhaps unlicensed, for communicating information, the latency in communicating information, etc.), not mentioned in this chapter are assumed to be handled appropriately.

5.2.1 Global Spatio-temporal Sensing with SAL

As a proximity based localization algorithm, SAL may provide a less complex and perhaps practical solution to the spectrum sensing problem. In this section, we explain how SAL can be utilized in global sensing of the spatio-temporal spectrum holes.

In SAL, a given number of monitoring stations with known locations are placed in a localization area A. When a primary transmitter is active, MoSs determine the subarea in which the activity is detected. These subareas are called *active* subareas. Figure 5.1 shows three active primary transmitters and three secondary transmitters. Assuming that the secondary and primary networks use the same frequency channel, only two of the secondary users are allowed to transmit due to the locations of active primary users localized in three subareas. The transmission ranges are shown as red circles.

To investigate the performance of SAL in spatio-temporal sensing, we assume the location of primary transmitters are estimated by SAL and primary receivers are at the worst case locations. A primary receiver is located at a worst case location if it is located on the edge of the transmission range of the primary transmitter. This is called worst case location, because this is the case in which primary receiver is the closest to a potential secondary transmitter. In Figure 5.2, the worst case position (WCP) for a primary receiver is shown.

In order to determine a spectrum hole so that the primary constraint is satisfied, we assume the localization system knows the transmission power of primary transmitters, P_t, and receiver sensitivity of primary receivers is at least R_s. Then, the average (ideal) transmission range, d_t, of a primary transmitter can be calculated from a typical path-loss model from,

$$R_s = P_t - PL(0) - 10\alpha log_{10}(d_t) - 10\beta log_{10}(f), \tag{5.1}$$

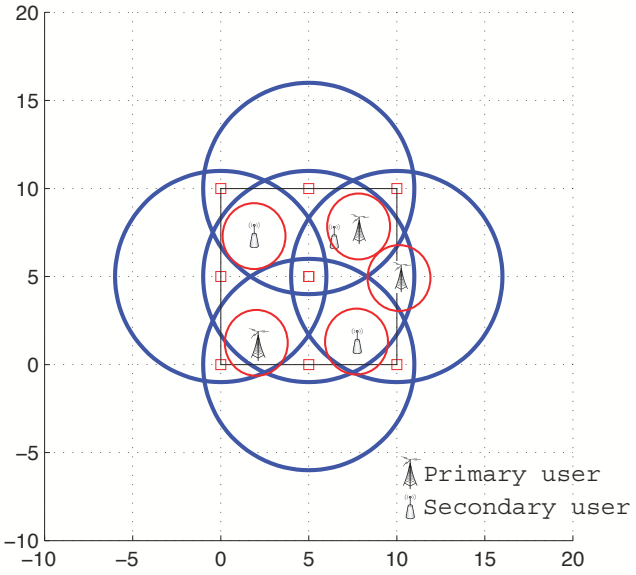

Figure 5.1: SAL usage in CWN

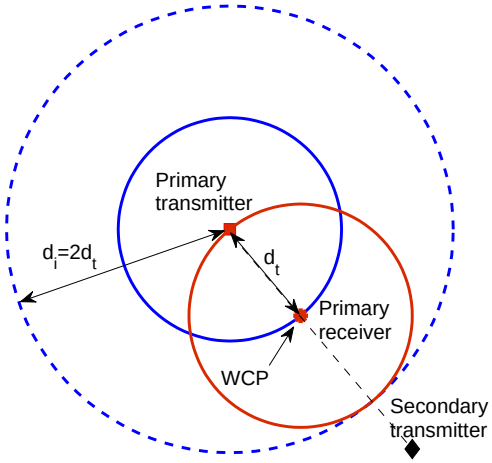

Figure 5.2: Worst case position for primary receiver

where f is frequency on which the signal is transmitted, and α, β and $PL(0)$ can be set to simulate the channel according to the environment.

With power P_t and sensitivity R_s, a primary transmitter can transmit up to a range of d_t, and a primary user can receive the signal at a maximum range of d_t. In the worst case, the interference range that is subject to interference from secondary users may be defined as the circle centered at the primary transmitter with $d_i = 2 \times d_t$ radius. The interference range and worst case positions are shown in Figure 5.2.

When a primary transmitter's location is estimated by SAL, the estimated location is the center of mass of the active subarea. Depending on the subarea size, the primary transmitter can be very close to the estimated location coordinates; however, it can also be at the edge of the active subarea. In order to satisfy the primary constraints protecting primary users from interference, the active subarea size must be extended by $d_i = 2 \times d_t$ in all directions. According to this very conservative approach, secondary users are not allowed to transmit if they are in or at the edge of this extended subarea. We also assume secondary transmitters use a power less than P_t to transmit their signals.

5.3 EXAMINATION OF THE POTENTIAL OF SAL FOR SECONDARY SPECTRUM USE

5.3.1 Evaluation of Efficiency of SAL in Detecting Primary Transmitters

The efficiency of SAL in determining spectrum holes depends on how close the estimated active area is to the real active area that a secondary user should not be not allowed to utilize for opportunistic spectrum access.

For a given monitoring range, and a number of monitoring stations, SAL determines subareas and regions as explained in Chapter 3. When the primary receiver locations are unknown and they are assumed to be placed in the worst case position, the interference range is calculated as $d_i = 2d_t$, where d_t is the transmission range of primary transmitters. When SAL detects a primary transmitter in a region R_j, it does not know where in the region it

is actually located. Although the estimated location is the center of mass of the region, the primary transmitter may also be located at the edge of the region. Therefore, in order to minimize the possible interference to primary receivers, SAL enlarges the region area by d_i. We denote the new extended region area as R'_j. Then, in the worst case, the actual occupied area is πd_i^2, and estimated occupied area is R'_j. We define the efficiency of spectrum hole detection in region R_j as

$$\eta_j(d, d_i) = \frac{A - \pi d_i^2}{A - R'_j} \tag{5.2}$$

Assuming all locations in A are equally likely for a primary transmitter, the expected efficiency in detecting spectrum holes in area A (η_{sh}) is calculated as,

$$\begin{aligned} \eta_{sh} &= \sum_{j=1}^{|R|} k_j \ \textbf{Prob}(\text{Primary transmitter is detected in region } R_j) \times \eta_j(d, d_i) \\ \eta_{sh} &= \sum_{j=1}^{|R|} \frac{k_j R_j}{A} \times \frac{A - \pi d_i^2}{A - R'_j}, \end{aligned} \tag{5.3}$$

where k_j is the number of regions R_j in A. We emphasize that in the best case $\eta_j(d, d_i) = 1$ as this will reflect the condition that the estimated occupied region is exactly the same size as the real active region.

In the following, we will use localization scenario $< 4, \mathbf{G}, A >$ as an example to show the performance of SAL in detecting spectrum holes.

To understand the performance, we start with computing the *sizes* of the subareas with $N = 4$. In Chapter 3, we showed that $\sqrt{2}g/2 < d < g$ is the optimal interval for monitoring ranges for minimum localization error. In this interval, 13 distinct subareas are formed and they can be classified into 4 types of regions, $R_i,\ i = 1, 2, 3, 4$ as shown again in Figure 5.3. Let C_j be the monitoring area of MoS_j, and the centers of C_1, C_2 and C_4 be located at $(0, L)$, (L, L) and $(L, 0)$, respectively. The calculation of sizes of regions R_i based on the

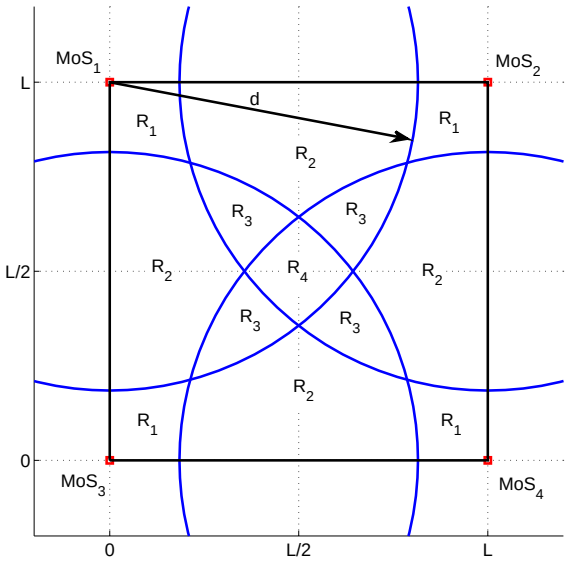

Figure 5.3: Visualization of regions in A when $\sqrt{2}L/2 < d < L$

intersection areas $B_1(d,g) = C_1 \cap C_2$ and $B_2(d,g) = C_1 \cap C_4$ as functions of d and g are given as:

$$B_1(d,g) = 2(R_2 + 2R_3 + R_4) = 4\left[\frac{\pi d^2}{2\pi/\arccos(g/2d)} - \frac{\sqrt{d^2 - g^2/4}}{4}\right] \tag{5.4}$$

$$B_2(d,g) = 2R_3 + R_4 = 4\left[\frac{\pi d^2}{2\pi/\arccos(\sqrt{2}g/2d)} - \frac{\sqrt{d^2 - g^2/2}}{4}\right] \tag{5.5}$$

Other relations between R_i and L can be derived as follows:

$$4(R_1 + R_2 + R_3) + R_4 = L^2 \tag{5.6}$$

$$R_1 + 2R_2 + 3R_3 + R_4 = \frac{\pi d^2}{4}. \tag{5.7}$$

Then the areas of regions $R_i,\ i = 1, 2, 3, 4$ can be calculated as,

$$R_1 = \frac{B_2(d, g) + g^2}{2} + \frac{\pi d^2}{4} \tag{5.8}$$

$$R_2 = \frac{B_1(d, g)}{2} - B_2(d, g) \tag{5.9}$$

$$R_3 = \frac{B_2 - g^2 + \pi d^2 - 2B_1(d, g)}{2} \tag{5.10}$$

$$R_4 = g^2 - \pi d^2 + 2B_1(d, g) \tag{5.11}$$

In order to calculate the extended regions R'_j, we need to decide how to modify the monitoring ranges of MoSs. Figure 5.4 shows how the circles must be modified when an activity is detected in subarea A_5 which can be classified as an R_2. The actual monitoring ranges are shown with blue circles and the red circles denote the modified circles. In order to calculate the extended region's area R'_2, the radius of C_2 and C_4 must be increased, and the radius of C_1 and C_3 must be decreased by d_i. Although there exist other correct sets of modifications, Table 5.1 shows one of these sets to calculate the area of the extended regions. d_m denotes the monitoring range of MoS_m, $m = 1, 2, 3, 4$.

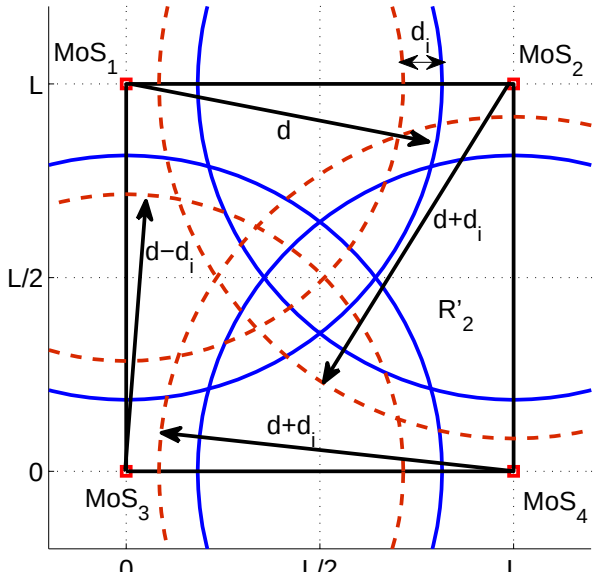

Figure 5.4: Visualization of extended region R'_2

Table 5.1: MoS ranges used in calculating enlarged regions with 4 MoSs

Regions, R'_j	d_1	d_2	d_3	d_4	k_j
R'_1	$d+d_i$	$d-d_i$	$d-d_i$	$d+d_i$	4
R'_2	$d-d_i$	$d+d_i$	$d-d_i$	$d+d_i$	4
R'_3	$d-d_i$	$d+d_i$	$d+d_i$	$d-d_i$	4
R'_4	$d+d_i$	$d+d_i$	$d+d_i$	$d+d_i$	1

The calculation of extended regions is not as straight forward as the R_i calculations shown previously because the circles that are overlapping have different radii in this case. In the following we will show how R'_is can be calculated for scenario $< 4, \mathbf{G}, A >$ and when $\sqrt{2}g/2 < d < g$.

First, we need the size of the intersection areas of two and three circles with different radii. They are given as,

$$C_1(d_1) \cap C_2(d_2) = \left(\frac{d_2^2\theta_2}{2} - x_2\ y\right) + \left(\frac{d_1^2\theta_1}{2} - x_1\ y\right) \tag{5.12}$$

$$\theta_j = \arccos\left(\frac{2y^2-1}{d_j^2}\right),\ j = 1,2$$

$$(x_1, x_2) = \left(\frac{g^2+d_1^2-d_2^2}{2g},\ \frac{g^2-d_1^2+d_2^2}{2g}\right),\ x_1 + x_2 = g$$

$$y = d_j^2 - x_j^2,\ j = 1,2,$$

$$C_2(d_2) \cap C_3(d_3) \cap C_4(d_4) = \sum_{j=2,3,4}\left(\frac{d_j^2\ \theta_j}{2} - T_j\right) + \frac{a_4\sqrt{a_2^2 - \frac{a_4^2+a_2^2}{2a_4}}}{2} \tag{5.13}$$

$$T_j = \frac{a_j\sqrt{4d_j^2 - a_j^2}}{4}$$

$$\theta_j = \arccos(\frac{a_j^2 - 2d_j^2}{2d_j^2})$$

$$a_2 = |c_3\ c_4|,\ a_3 = |c_2\ c_4|,\ a_4 = |c_2\ c_3|$$

where θ is in radians, and c_2, c_3 and c_4 are the intersection points shown in Figure 5.5.

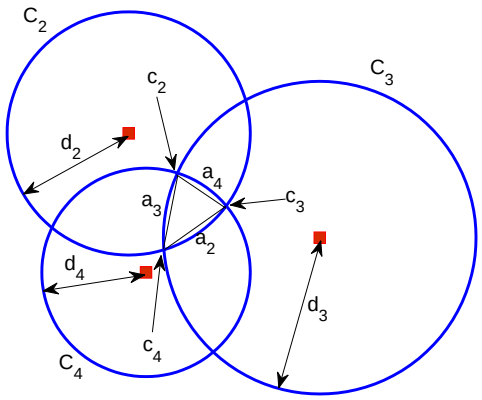

Figure 5.5: Intersection points of three circles C_2, C_3 and C_4

Then, the areas of extended regions $R'_i,\ i = 1, 2, 3, 4$ can be calculated as,

$$R'_1 = \frac{R'_4 + g^2}{4} - (R'_2 + R'_3) \tag{5.14}$$

$$R'_2 = \frac{B_{1^+2^+}}{2} - B_{2^+3^-} + B_{2^+3^-4^-} - B_{1^+2^+4^-} \tag{5.15}$$

$$R'_3 = B_{2^+3^+} - B_{2^+3^+4^-} \tag{5.16}$$

$$R'_4 = g^2 - \pi(d + d_i)^2 + 2B_{1^+2^+} \tag{5.17}$$

where the intersections are given as,

$$B_{1^+2^+} = C_1(d + d_i) \cap C_2(d + d_i) \tag{5.18}$$

$$B_{2^+3^-} = C_2(d + d_i) \cap C_3(d - d_i) \tag{5.19}$$

$$B_{2^+3^+} = C_2(d + d_i) \cap C_3(d + d_i) \tag{5.20}$$

$$B_{2^+3^-4^-} = C_2(d + d_i) \cap C_3(d - d_i) \cap C_4(d - d_i) \tag{5.21}$$

$$B_{2^+3^+4^-} = C_2(d + d_i) \cap C_3(d + d_i) \cap C_4(d - d_i) \tag{5.22}$$

$$B_{1^+2^+4^-} = C_1(d + d_i) \cap C_2(d + d_i) \cap C_4(d - d_i) \tag{5.23}$$

We demonstrate the spectrum hole detecting efficiency with scenario $< 4, \mathbf{G}, A >$ in Figure 5.6 with respect to the monitoring range and interference range. Since it is not possible to estimate the exact real occupied region, the efficiency is always larger than 1, which is the best possible value. The red line in this figure shows the conditions which minimize η_{sh} for a given interference range, meaning that these make it approach unity. Figure 5.7 shows the spectrum hole efficiency and best efficiency conditions with scenario $< 9, \mathbf{G}, A >$. As expected, in both figures, the observation is that as interference range increases the efficiency moves further away from unity. This is due to larger areas of estimated occupied regions with increasing interference range. Another observation is that there is an optimum monitoring range minimizing η_{sh} for a given interference range.

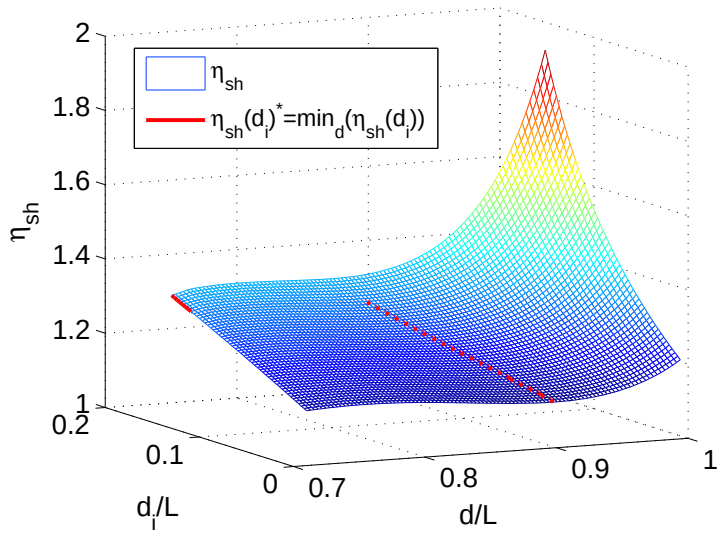

Figure 5.6: $N = 4$ MoSs, η_{sh}

The best efficiency $\eta^*_{sh} = min_d\left[\eta_{sh}(d_i)\right]$ with $< 4, \mathbf{G}, A >$ and $< 9, \mathbf{G}, A >$ is shown in Figure 5.8 as a function of the interference range. With 9 MoSs in grid placement, 7 different regions (calculated and shown in Chapter 3) exist, and they are smaller than the regions formed with 4 MoSs in grid placement. Therefore, the efficiency with 9 MoSs is closer to unity than with 4 MoSs. Figure 5.9 shows the best monitoring ranges (d^*) for a given interference range (d_i), and $d^* + d_i$. We observe that as d_i is increasing, best monitoring range

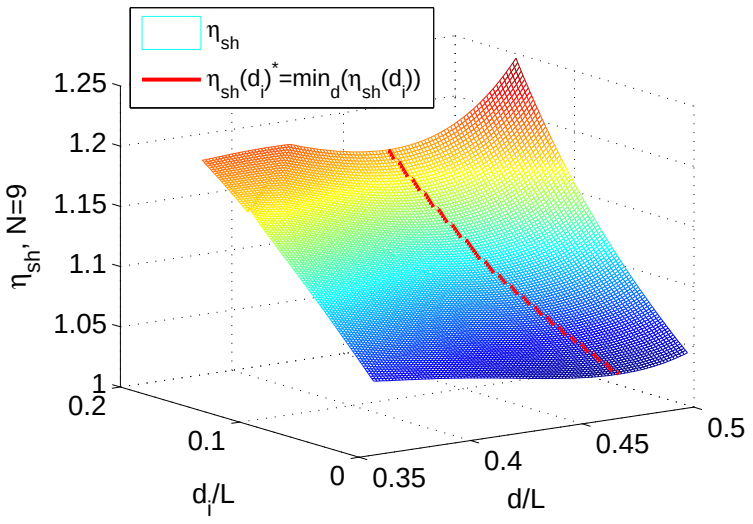

Figure 5.7: $N = 9$ MoSs, η_{sh}

decreases. In case of 4 MoSs, when $d_i > 0.12L$, the best monitoring range drops sharply to its minimum value ($0.707L$, which is needed to cover the whole area $A = L \times L$), although there is a negligible difference in the minimum η^*, if the slope had continued as it was when $d_i < 0.12L$.

5.3.2 Effect of Frequency Offset on Location Estimation

Although the primary transmitter signal is detected, it is possible that the monitoring ranges are optimized for a frequency f_{LS}, and so the localization accuracy of transmitters using frequencies other than f_{LS} may be different. In this section, we will explore the effect of the difference between the localization system frequency and primary transmitter frequency on the accuracy of estimated location of primary transmitters.

5.3.2.1 Free Space Path Loss Model According to the free space path loss model, the path loss in dB is calculated as,

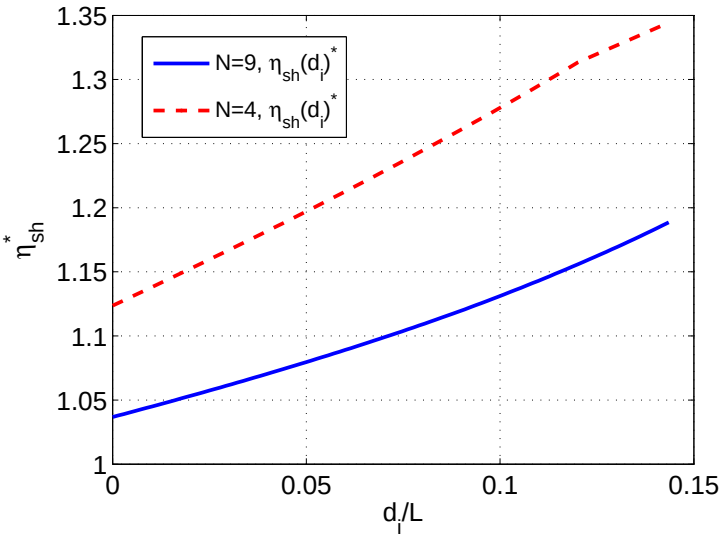

Figure 5.8: Spectrum hole detection efficiency, η^*_{sh}

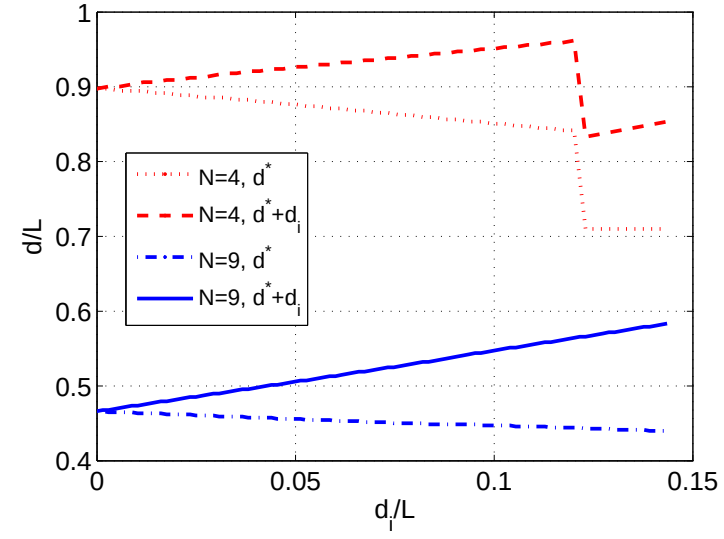

Figure 5.9: Best monitoring range at η^*_{sh}

$$PL(dB) = 32.4 + 20log_{10}(d) + 20log_{10}(f), \tag{5.24}$$

where d is the distance between the transmitter and the receiver in kilometers and f is the frequency of signal in MHz. The free space path loss model can be used if the primary transmitter height and monitoring station heights are larger than the building height. In this case, building heights and street width may not significantly impact the signal propagation from transmitter to the receivers – MoSs.

In Figure 5.10, we assume that the localization system has been optimized with the best monitoring ranges assuming that the signal frequency of primary transmitters is f_{LS} = 900MHz under ideal channel conditions. According to this figure, since the optimization is done according to f_{LS} = 900MHz, the accuracy at this frequency is the best. The accuracy gets worse while the primary transmitter frequency is getting smaller, while the availability is kept at 100%. However, the availability decreases significantly when the primary transmitter frequency is increasing. The reason is the change in the MoS's coverage while increasing/decreasing the frequency. Because of low availability at these frequencies, the localization error seems to be lower than the value achieved at f_{LS}. The efficiency achieved at best accuracy conditions is shown in the bottom most figure. According to this, the highest efficiency is achieved at f_{LS} = 900MHz, and it starts decreasing towards both lower and higher frequency sides. At f = 850MHz, the efficiency is similar to f_{LS} due to similar accuracy and availability conditions. By employing CA, we can provide better efficiency for frequencies smaller than f_{LS}. SAL+CA and SAL performs similarly under ideal channel conditions.

Figure 5.11 shows the best accuracy conditions, and their effects on availability and efficiency when f_{LS} = 1900MHz under ideal conditions. For $f \geq f_{LS}$, the accuracy of SAL is better than CA. The effect on availability is not significant; therefore, the efficiency is higher compared to CA. However, for $1200MHz < f < f_{LS}$, the localization error with SAL is larger than the error with CA, the availability is 100%. And for $f < 900$, the performance of CA and SAL are similar.

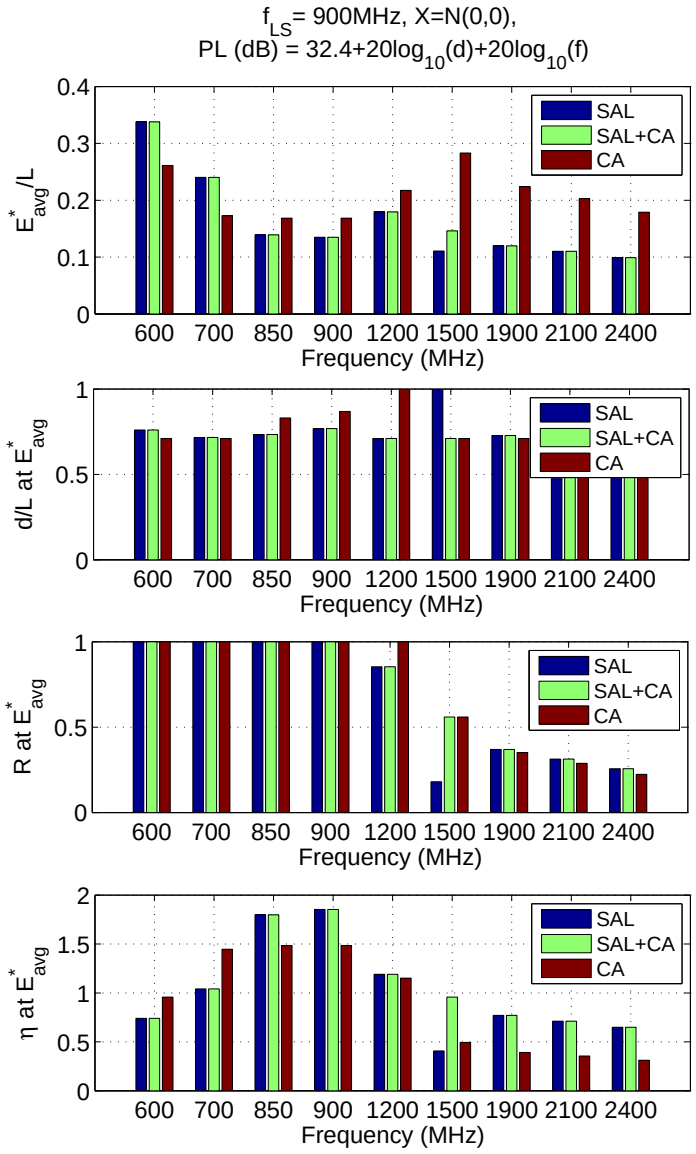

Figure 5.10: Performance when optimized for LS, $N = 4$ MoSs, free space path loss model, $f_{LS} = 900$MHz

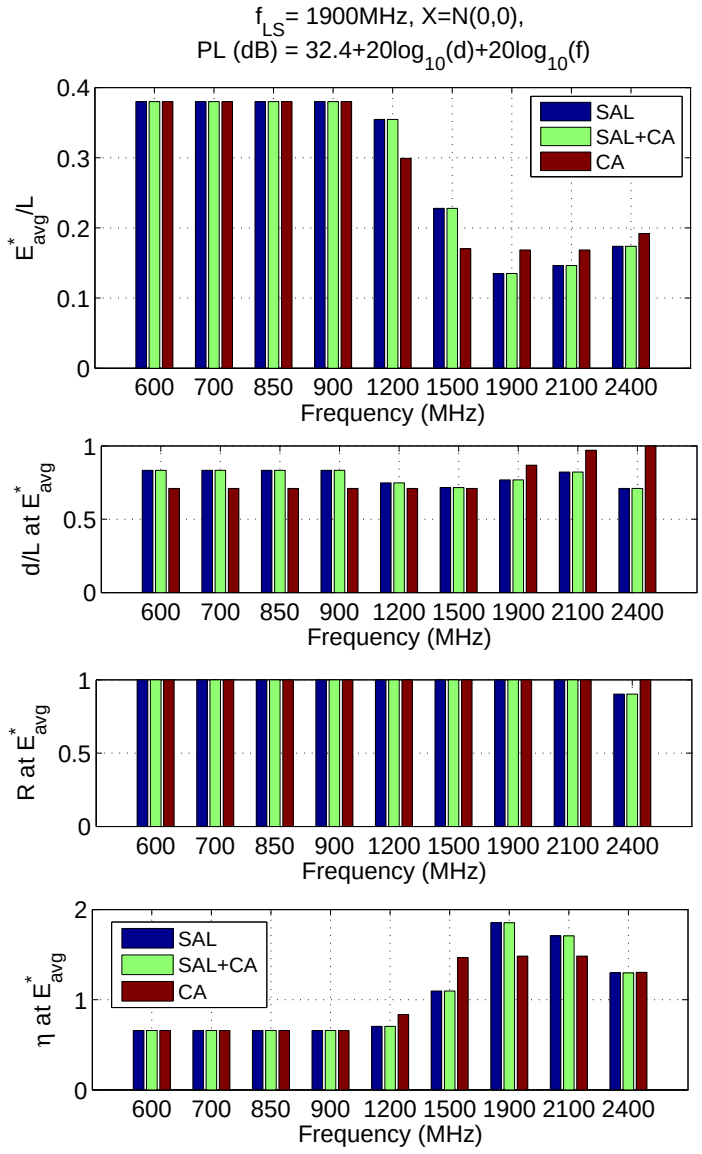

Figure 5.11: Performance when optimized for LS, $N = 4$ MoSs, free space path loss model, $f_{LS} = 1900$MHz

5.3.2.2 Near Building Height Path Loss Model

A path loss model that incorporates parameters such as the building height is suitable for the case when primary transmitters are at the same height as the buildings surrounding the area, and MoSs are at a lower height than the buildings. Therefore, the building height, monitoring station antenna height and street width must be considered in the signal propagation model. According to one such model [85], the path loss in dB is calculated from,

$$PL(dB) = 42.12 + 40log_{10}(d) + 30log_{10}(f), \tag{5.25}$$

where d is the distance between the transmitter and the receiver in kilometers and f is the frequency of signal in MHz when the building height is $h_{build} = 12$m, primary transmitter height is $h_{pt} = 12$m, and monitoring station height is $h_m = 7$m, and street width is $w = 30$m.

We repeat the simulations above with this building height path loss model. Figure 5.12 shows the best accuracy conditions when $f_{LS} = 900$MHz under ideal channel conditions. Figures 5.13 shows the results of a similar simulation with $f_{LS} = 1900$MHz under ideal channel conditions. The observations are similar to the ones assuming free space path loss model, although the effect of distance and frequency on path loss is more with the building height influenced path loss model.

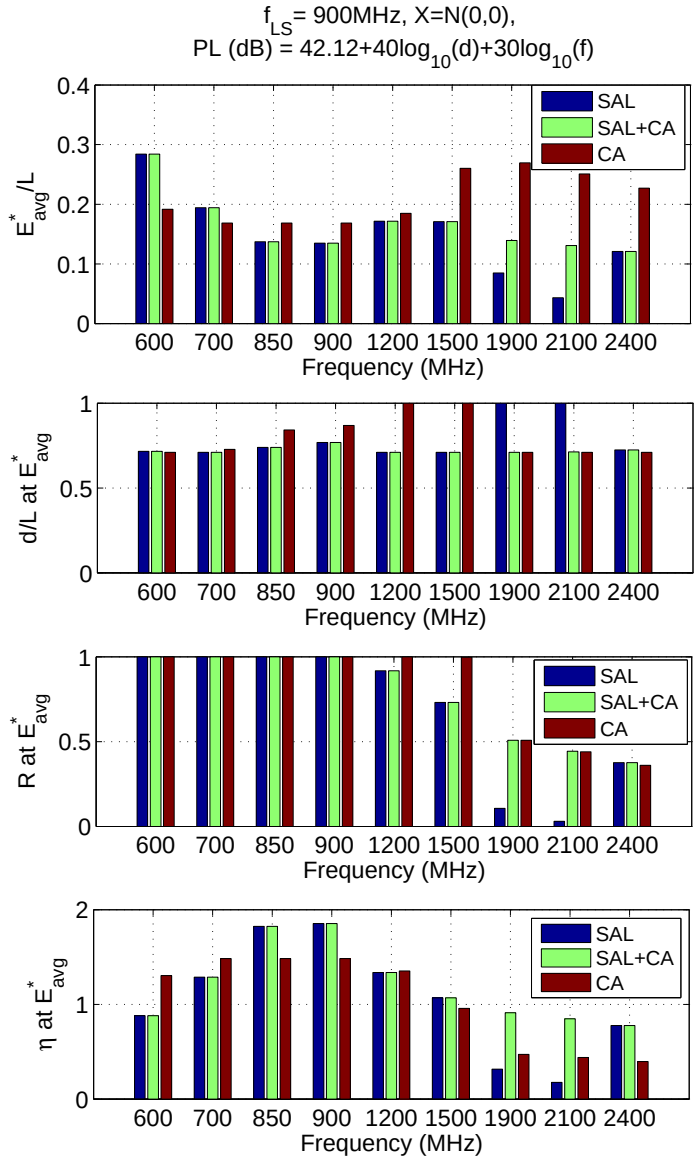

Figure 5.12: Performance when optimized for LS, $N = 4$ MoSs, building height, $f_{LS} =$ 900MHz

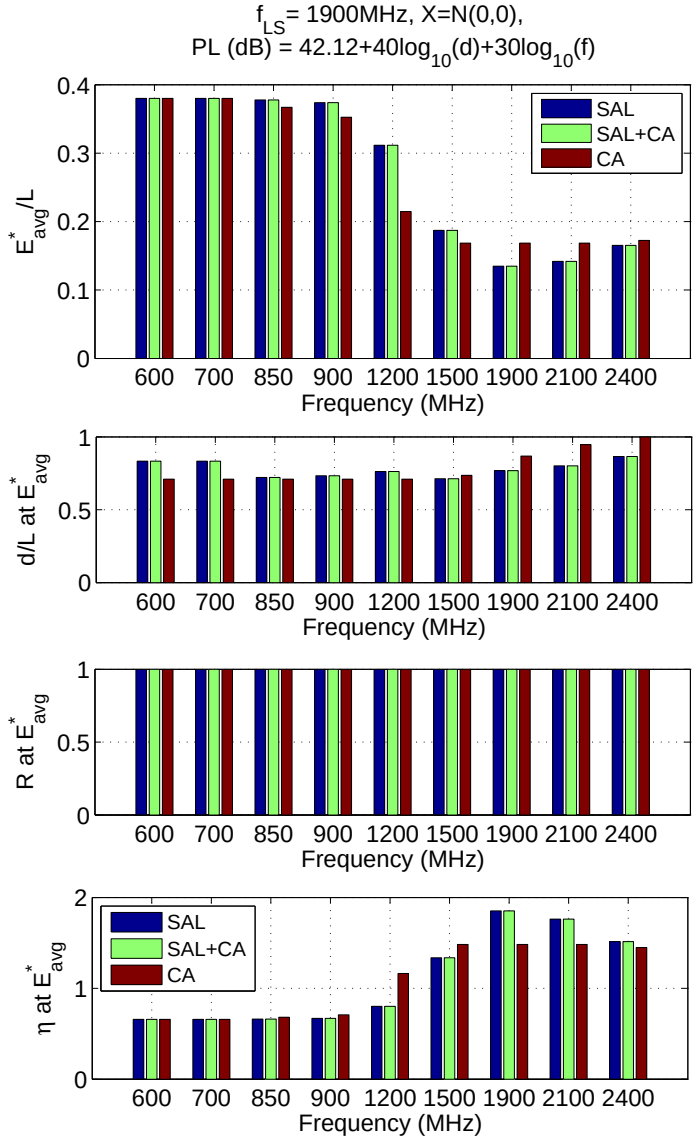

Figure 5.13: Performance when optimized for LS, $N = 4$ MoSs, building height, $f_{LS} =$ 1900MHz

5.3.3 Effect of Frequency Offset on Spectrum Hole Detection

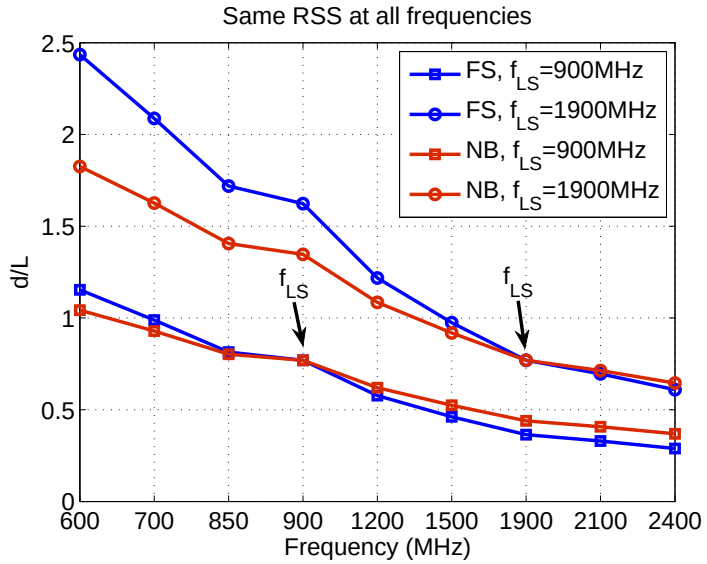

Figure 5.14: Change of monitoring range with frequency offset

The frequency offset between primary transmitters and localization system causes differences between actual subareas and subareas known by SAL. Figure 5.14 shows how the actual monitoring ranges change under various frequency offsets with free space (FS) and near building (NB) height path loss models. For example, when the frequency of the localization system is 900MHz, the best monitoring range is given as $d = 0.77L$; however, the same transmission power and same receiver sensitivity can generate a monitoring range of $1.15L$ at 600MHz, and $0.28L$ at 2.4GHz with FS path loss model. Due to these differences, the actual subareas may shift, shrink or expand compared to subareas known by SAL.

Figures 5.15 and 5.16 show the impact of frequency offset on localization scenario $< 4, \mathbf{G}, A >$, when $f_{LS} = 900MHz$, and the primary user frequency is 700MHz and 1.2GHz, respectively. When $f < f_{LS}$ in Figure 5.15, we observe that R_1 shrinks, R_2 shrinks on one dimension, and expands on other dimension, R_3 expands and shifts and R_4 expands. And when $f > f_{LS}$ in Figure 5.16, we observe that R_1 expands, R_2 shrinks in one dimension, and expands in other dimensions, R_3 shrinks and shifts and R_4 shrinks.

After the detection of active subarea, if the frequency information is not utilized in

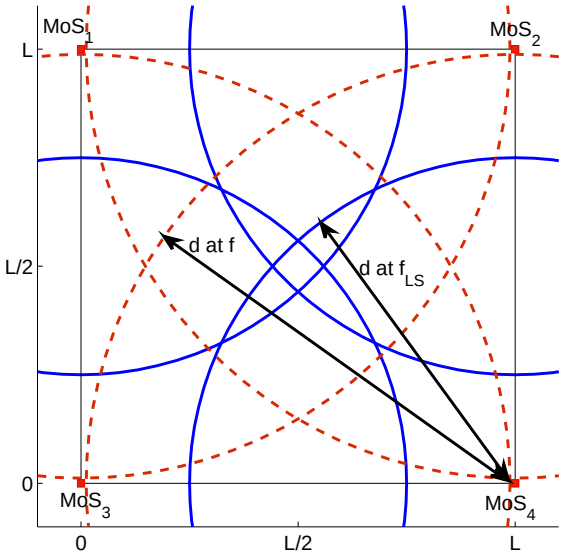

Figure 5.15: Visualization of regions when $f_{LS} = 900\text{MHz}$ and $f = 700\text{MHz}$

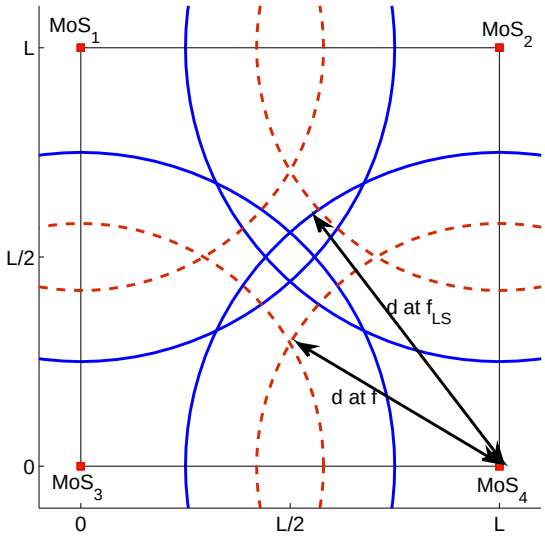

Figure 5.16: Visualization of regions when $f_{LS} = 900\text{MHz}$ and $f = 1.2\text{GHz}$

estimation of new regions, these differences will reflect as differences in extended regions, and in turn the accuracy of spectrum hole detection will be impacted negatively. In this case, when regions expand primary users may be vulnerable to interference from secondary users, and when regions shrink, underutilization of available spectrum occurs. In order to get better accuracy on detection of spectrum holes, the primary user frequency information must be used to estimate the extended regions. Although the detection of active subarea is done with the monitoring ranges that are calculated with f_{LS}, the primary user frequency can be used to estimate (conservatively), the spectrum holes.

5.4 DISCUSSION ON POSSIBLE RESEARCH DIRECTIONS

The performance of SAL in detecting spectrum holes can be improved by available information such as primary receiver locations, protocols used in communication in primary and secondary networks and secondary transmitter locations, etc. In this section, we will discuss how additional information on such parameters may affect the spectrum hole detection.

SAL based spectrum hole detection method proposed in previous sections only use monitoring stations that are deployed at fixed locations. Employing the secondary network in detection of spectrum holes could help satisfy the primary and secondary constraints, and also narrow down the subarea size at the vicinity of secondary nodes, so that a more accurate estimation of spectrum holes is possible. Towards this goal, an approach may be to use secondary receivers for local sensing while using fixed monitoring stations for global sensing. With SAL, secondary nodes can be notified of the locations of primary transmitters. When needed, they can perform local sensing to guarantee interference free transmission, but need not share this information with other secondary transmitters. Another approach may be to utilize secondary nodes with known locations to help monitoring stations in detecting primary transmitters in subareas during the spatio-temporal sensing phase. While one drawback of this approach may be that the exact location of secondary nodes must be known, it allows us to determine an upper bound on spatio-temporal sensing performance and throughput of primary users, and spectrum utilization by secondary nodes. An ex-

amination of this scenario would be similar to random placement of MoSs combined with SAL.

Another research direction to take could be to employ secondary nodes as monitoring stations in SAL instead of deploying additional fixed monitoring stations for location estimation of primary users. This approach may reduce the overhead in transferring topology information to secondary users; however, there are a number of challenges in implementation of this approach. A secondary node must know the location of itself, be loosely synchronized with the localization system, and communicate this information to the localization center. The number and the spatial distribution and movement characteristics of the secondary nodes may have a large impact on the the spatio-temporal sensing accuracy. The amount of clustering and neighborhood properties and environmental characteristics would affect the performance. The trade of between using only secondary nodes or only MoSs in SAL must be investigated to find the optimum number of MoSs for global sensing and optimum number of secondary nodes with known locations for local sensing.

In opportunistic access configuration phase, secondary users decide whether and how to exploit spectrum opportunities, what transmission power to use and how to share opportunities among secondary users. Therefore, in this phase the objective is to determine the transmission range of secondary users while satisfying the primary and secondary constraints at the same time. Another objective is to provide fair utilization of spectrum holes by secondary transmitters. Fair utilization can be achieved by balancing the transmission ranges of secondary transmitters as much as possible while maximizing the utilization of spectrum holes. SAL may be useful in these aspects as well.

In previous sections, we assumed primary receiver locations are not known and are not estimated. They are assumed to be located at the worst case position. If the primary network only consists of transmitters which are broadcasting, and not exchanging signals with receivers, then, estimating locations of primary receivers is more important than estimating the locations of primary transmitters because the interference range will only be around the receivers, transmitter will not get affected by interference. If we assume the primary receivers are at the worst case positions, this causes a conservative approach and the estimated spectrum hole may be actually smaller than the real spectrum hole. When primary

receiver locations are known, the actual occupied area can be estimated with more accuracy and therefore the spectrum hole detection performance improves. If the primary network uses a protocol such as CSMA-CA where the nodes first listen to the channel, and do not transmit if there is a transmission by a secondary node going on, and back off for collision avoidance, then secondary transmissions must not interfere with the transmitter either. In this case, primary transmitter locations must be estimated because primary network is the owner of the spectrum, and has the priority to use it when needed.

As a proximity based localization algorithm Sub-Area Localization can provide an advantage in terms of cost and responsiveness due to the low complexity of algorithm. However, as discussed in previous chapters, SAL availability decreases significantly under severe shadowing. In this case, using SAL+CA will be a better approach in detecting primary transmitters.

6.0 CONCLUSIONS AND FUTURE WORK

Localization with wireless networks with many of its applications is a very interesting research area that it has been continuing to draw the attention of researchers for many years. Localization approaches can be generally classified as fine grained and coarse-grained depending on the granularity of information inferred by the communication between mobile devices and reference points. We surveyed the literature extensively in order to understand the characteristics, drawbacks, and strengths of many WiFi RSS based localization algorithms that have been proposed previously. Although it is not straightforward to compare these systems, we have developed metrics that use the data presented in the respective research related to these systems to compare them quantitatively.

In this dissertation, the primary interest was in the potentials of proximity based localization which is a coarse grained approach using connectivity information to reference nodes. This approach may have worse accuracy; however, it does not suffer from multi-path and hardware related problems or laborious fingerprint collection associated with fine grained localization. This is maybe the most attractive feature of proximity based localization.

In order to study the potential of proximity based localization, we proposed Sub-Area Localization (SAL) which is a 2-dimensional proximity-based technique. SAL utilizes the joint monitoring regions –subareas– of multiple monitoring stations. We showed through analytical and simulation results that there exists an optimal coverage (range) for best localization accuracy with SAL given a localization scenario with a specific number of monitoring stations covering the network area. Our analyses and simulations are enriched with performance metrics such as availability and efficiency that help make a fair and comprehensive comparison between different localization scenarios and algorithms. With optimized monitoring ranges, we showed that SAL can provide at least 30 cm better accuracy compared to

the previously proposed Centroid Algorithm in a localization area of size $100m^2$.

Our research in 2-dimensional proximity based localization revealed that increasing the number of monitoring stations can certainly reduce the localization error and monitoring range and this follows an exponential drop in error however incurring increasing cost. Multi-floor buildings could provide an opportunity to improve localization accuracy without increasing the number of the infrastructure entities. Consequently, we examined an interesting aspect namely whether combining the readings from monitoring stations deployed in different floors is possible for better coverage and also better accuracy in each floor of a multi-floor building. We demonstrated that SAL can be used in multi-floor buildings, and we call this approach Multi floor Sub-Area Localization (MSAL). We investigated the performance of MSAL in two parts: when floor determination is externally available, and when a floor determination algorithm is integrated with MSAL. In both cases, we investigated the effect of antenna model, floor height, floor repetition scheme (for placing monitoring stations) and the variance in RSS due to shadow fading on the efficiency of the system. Our simulation results indicate that using MoSs across floors can increase the deployment efficiency by up to 50%. Our multi-floor localization system provides within-floor localization accuracy comparable to that on a single floor with fewer monitoring stations. As an example, if we consider the Trump World Tower in New York, USA, it has a floor count of 72 and a per floor area of $35 \times 35m^2$ (assuming a square floor area) making a total area of 89,800 m^2 [86]. The approximate normalized floor height is given as 3.5m from the calculation provided by the Council on Tall Buildings and Urban Habitat [87]. With this floor height, using a 2 floor repetition scheme, the efficiency can be doubled when compared to single floor localization approach. Under ideal conditions, when best monitoring ranges are determined, isotropic and dipole antenna coverages lead to similar localization error performances using MSAL with various floor repetition schemes (keeping the total number of visible monitoring stations per floor similar to that on a single floor with SAL). The best ratio of monitoring ranges of MoSs on adjacent floors is between 0.5 and 1.5 depending on floor height and shadow fading. However, as RSS shadowing fluctuations are included, system designers can select the same monitoring range on all floors without significantly sacrificing accuracy.

Among many applications of localization, spectrum sensing for identifying spectrum holes

in cognitive wireless networks is an interesting and yet an emerging application. At the end of this dissertation, we worked on a case study with the objective of showing that SAL can provide reasonable global awareness in terms of active primary transmitter locations, and help secondary networks utilize the available spectrum in time-space and frequency. We determine the interference and monitoring range conditions under which the spectrum detection efficiency reaches the best level. We evaluate the effect of two different path loss models and frequency offset between the localization system and primary network. We further contribute to this emerging research area with a discussion on possible research directions.

6.1 FUTURE WORK

In this dissertation, although we have done an extensive simulation based evaluation of sub area localization, an evaluation with real deployments of monitoring stations in appropriate indoor and outdoor environments is missing. With real deployments the effect of channel and obstructions in various environments and the effect of monitoring station placement may be captured more accurately.

Our approach in SAL and MSAL requires (or assumes) that there must be only one transmission at a time in the subareas. However, in real life, this may or may not be true. As future work, this requirement must be understood better and perhaps removed which will allow a localization system to distinguish simultaneous transmissions from different MNs, and to be able to detect and localize different MNs separately, yet simultaneously. Clearly, when there are simultaneous transmissions, using only the existence of a signal energy may cause problems. Signals may add up together when they arrive at MoSs unless they are separable, and an ambiguity will occur because the localization system is not aware if a transmission has multiple sources. This may require a more sophisticated localization algorithm. One proposal may be to utilize RSS measurements and potentially MN IDs collected at MoSs in addition to the connectivity information.

In this dissertation, our focus was on showing the impact of different placement schemes

on the performance of sub area localization approach instead of optimizing the placement of monitoring stations. A more extensive study on placement schemes may provide a better idea on the best placement schemes for various types of environments, single floor and multi floor, etc. Application scenarios may exist where physical access to the localization environment is not available. One example is helping firefighters to know where they are in a building that is on fire. In this case, inside deployment of monitoring stations may not possible and they must be deployed outside the building as this is possible and much more practical. Another application scenario is when the detection of intruders in wireless networks is needed.

For tractable analysis we assumed that the mobile devices transmit with the same signal power although different types or brands of wireless devices can have different specifications on the transmission power. Transmission power is one of the parameters that the localization system utilizes to calculate the average monitoring range. When the transmit power of mobile device is larger than the power known to the localization system, although the mobile device is beyond range d, it may appear that it is in range; and this may cause increased localization error. In the case that monitoring stations can determine and differentiate between the mobile devices, and have a database of transmit powers of various devices, this problem may be addressed without a need to fine tune the proximity threshold. Another way is to utilize transmission power control mechanisms.

We have compared the accuracy performance with monitoring stations equipped with isotropic antennas to the performance with modified dipole antennas while assuming mobile devices are equipped with isotropic antennas. Using dipole antennas, and furthermore, MIMO antennas or directional antennas may also have a significant impact on the performance due to their effect on transmission range and/or transmission direction. With directional antennas especially, the localization system will require the direction of transmission as an additional input in order to calculate the location estimate.

One interesting problem to look at is the effect of multiple communication technologies available on mobile devices, such as WiFi on various frequencies, Bluetooth, cell phone. Exploiting multiple wireless communication technologies may help improve accuracy if the localization algorithm is sophisticated enough to take advantage of it. Different frequencies and different transmission powers used in different wireless signals may cause different

localization performance under different channel and environmental effects. The cost of infrastructure elements in different technologies varies, as also the communication range. For example RFID tags have a very small range and are inexpensive, whereas Wi-Fi or Bluetooth devices have larger range and are more expensive than RFID tags. An interesting research can focus on efficient usage of the combinations of various communication technologies depending on the environment, required accuracy, available budget, etc. Furthermore, one may not care about minimizing the cost when the required accuracy is satisfied; or one may not care about improving the accuracy as long as "a" location estimate is available for a large fraction of the localization area. Determining the trade off between these parameters can be another research direction.

We have used the term "cost" simply as the number of infrastructure entities deployed. As the above discussion clearly indicates, cost is a more complex function of type of device, application, environment, energy, and other factors that have not been included in this work. Accurately characterizing the cost based on various possibilities can be useful for system deployment.

Localization algorithms can be used as navigation and tracking mechanisms if the location estimation is calculated periodically with a given time interval in between estimates. Sub Area Localization algorithm can be used as a navigation system as its responsiveness only depends on how soon the monitoring stations can transfer their findings to the central server for location estimate calculation. Since the central server only maps the set of MoSs to a known subarea, the responsiveness of system can be expected to be satisfactory.

APPENDIX

DETAILED SIMULATION RESULTS FOR VARIOUS LOCALIZATION SCENARIOS

A.1 MSAL

A.1.1 Effect of Wall Attenuation Factor

Figure A1 shows the maximum efficiency with respect to σ and floor height for MSAL-iso and MSA-dpl in 2FRS scenario for 2 floor plans.

From Figures A2 and A3 show the behavior of maximum efficiency when $h = 0.05L$ and $h = 0.4L$, we can observe as floor height increases the maximum efficiency with isotropic and dipole antennas decreases and approaches to similar values. Also, at small h we see that dipole antennas can provide much better efficiency than isotropic antennas, whereas at large h, isotropic antennas can provide better results.

A.1.2 Effect of Random Placement

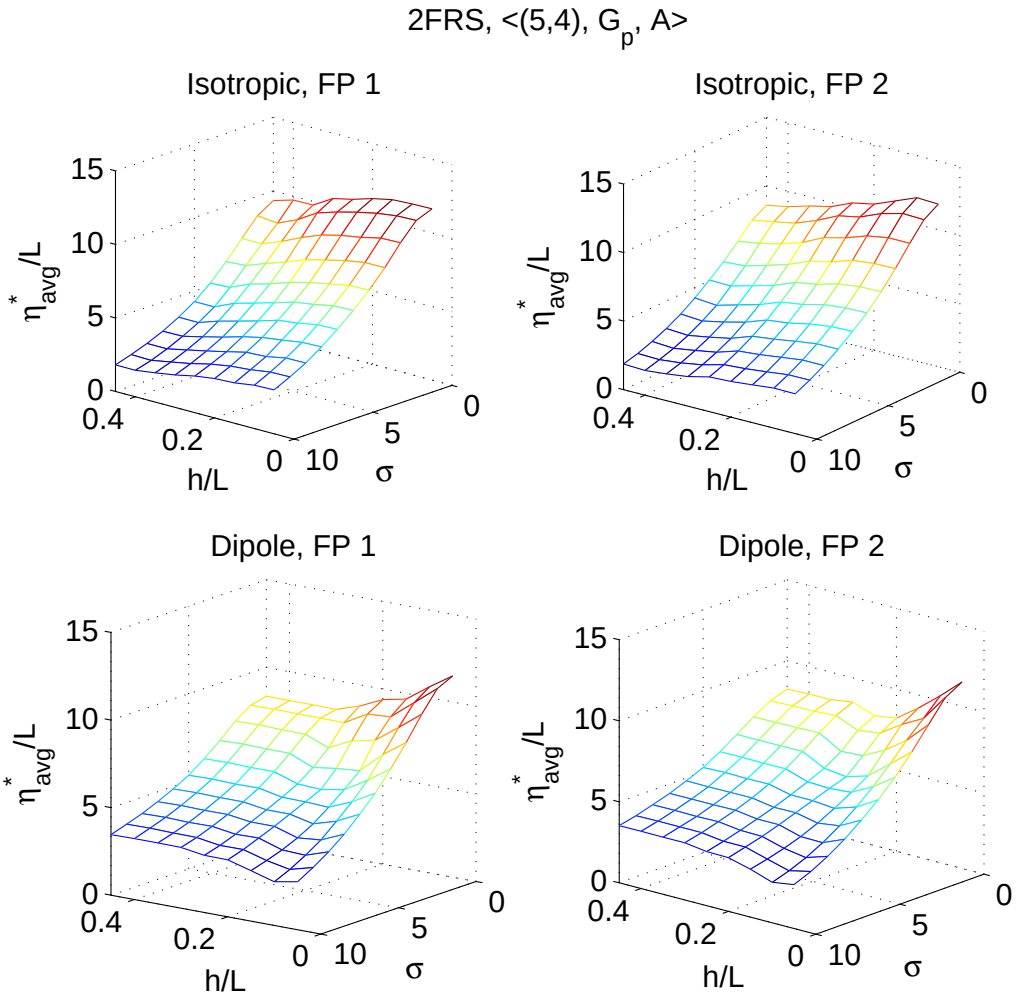

Figure A1: Maximum efficiency for different values of h and σ with 2FRS

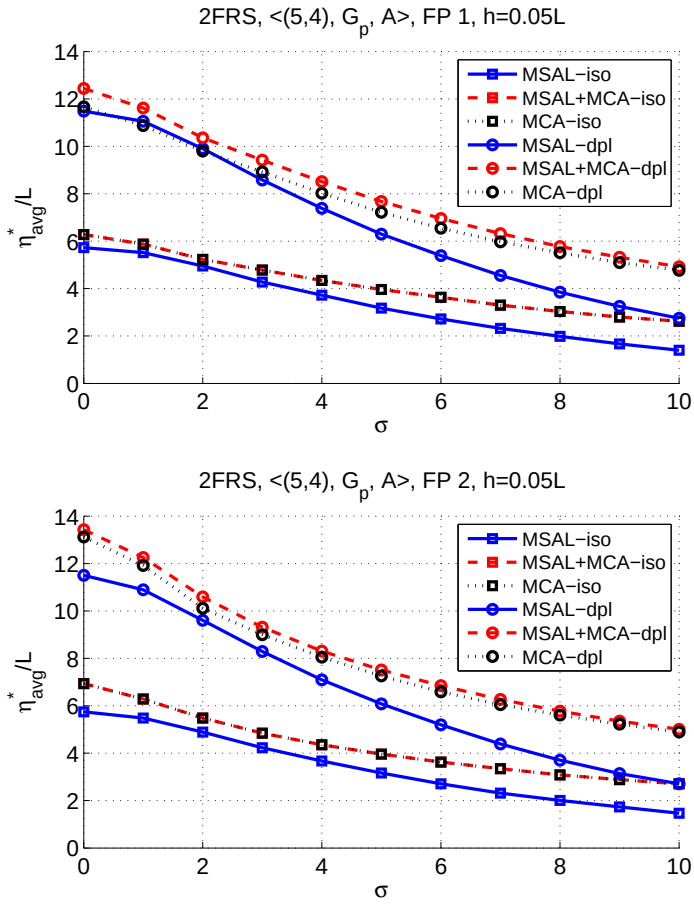

Figure A2: Maximum efficiency for different values of σ and $h = 0.05L$ with 2FRS

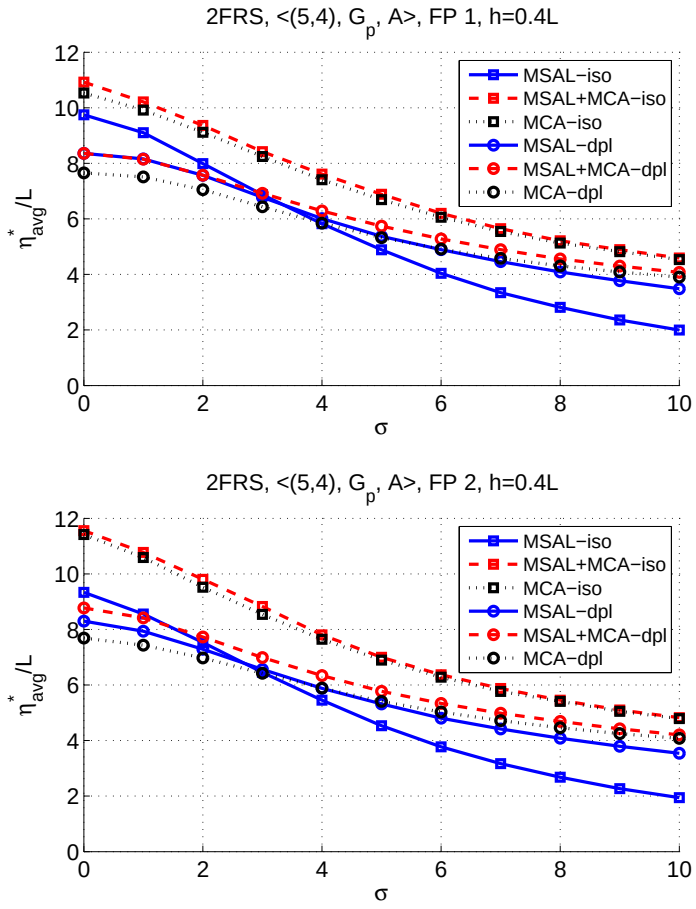

Figure A3: Maximum efficiency for different values of σ and $h = 0.4L$ with 2FRS

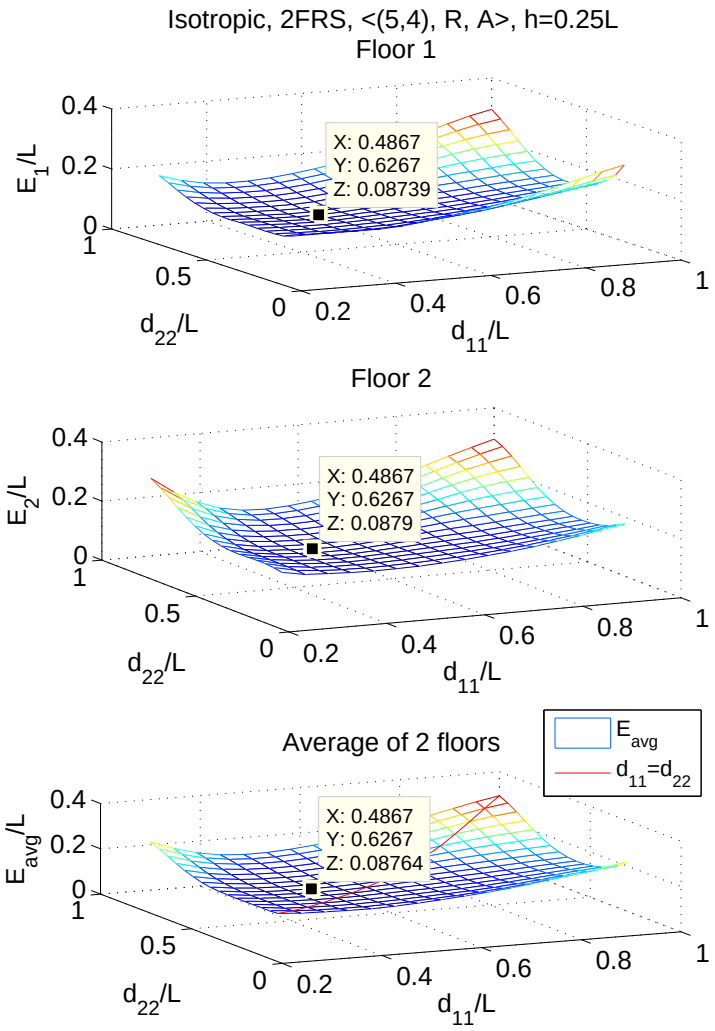

Figure A4: Localization error with 2FRS, $< (5,4), \mathbf{R}, A >$, $\sigma = 0$

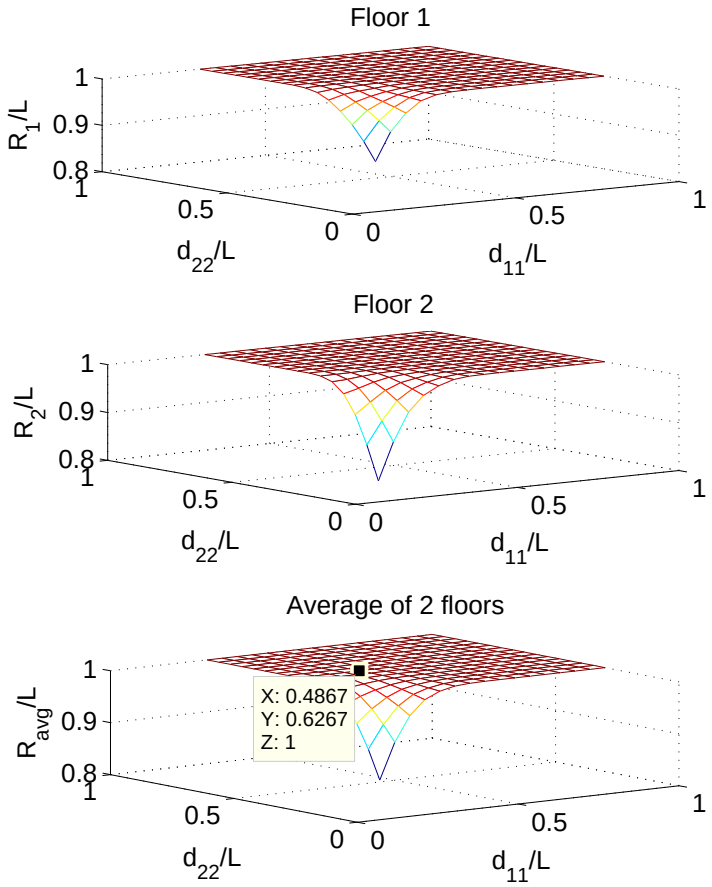

Figure A5: Availability with 2FRS, $< (5,4), \mathbf{R}, A >$, $\sigma = 0$

A.1.3 Summary of Results

In the following, we present the best accuracy and best efficiency values that can be achieved with MSAL, MSAL+MCA and MCA with various localization scenarios under ideal channel conditions and and two different levels of shadowing. We assumed $h = 0.25L$ which can correspond to 2.5m floor height in a building with floor are of $0 \times 10\ m^2$.

Tables A1 - A4 show the accuracy and efficiency performances of MSAL-iso with 2FRS under various localization scenarios and signal propagation conditions.

Tables A5 - A8 show the accuracy and efficiency performances of MSAL+MCA-iso with 2FRS under various localization scenarios and signal propagation conditions.

Tables A9 - A12 show the accuracy and efficiency performances of MCA-iso with 2FRS under various localization scenarios and signal propagation conditions.

Tables A13 - A16 show the accuracy and efficiency performances of MSAL+MCA-dpl with 2FRS under various localization scenarios and signal propagation conditions.

Tables A17 - A20 show the accuracy and efficiency performances of MCA-dpl with 2FRS under various localization scenarios and signal propagation conditions.

Table A1: MSAL accuracy with 2FRS in various localization scenarios

Scenario		$\sigma = 0$	$\sigma = 5$	$\sigma = 10$
		$h = 0.25L$	$h = 0.25L$	$h = 0.25L$
$<(5,4), \mathbf{G}_p, A>$	E^*_{avg}/L	0.081	0.172	0.251
	d^*_{ii}/L	(0.705 , 0.853)	(0.484 , 0.484)	(0.521 , 0.558)
$<(5,4), \mathbf{G}_p, A>$	E^*_{avg}/L	0.099	0.148	0.205
FP 1, $WAF = 3.1$dB	d^*/L	(0.484 , 0.447)	(0.484 , 0.484)	(0.447 , 0.447)
$<(5,4), \mathbf{G}_p, A>$	E^*_{avg}/L	0.088	0.146	0.205
FP 2, $WAF = 3.1$dB	d^*/L	(0.484 , 0.447)	(0.484 , 0.484)	(0.447 , 0.447)
$<(5,4), \mathbf{G}_p, A>$	E^*_{avg}/L	0.091	0.225	0.308
$FAF = 13$dB	d^*_{ii}/L	(0.889 , 0.889)	(0.521 , 0.521)	(0.779 , 0.558)
$<(5,4), \mathbf{R}, A, sh>$	E^*_{avg}/L	0.087	0.18	
	d^*_{ii}/L	(0.627, 0.487)	(0.44,0.44)	
$<(5,4), \mathbf{G}_p, A>$	E^*_{avg}/L	0.077	0.171	0.248
with MBFD	d^*_{ii}/L	(0.484 , 0.484)	(0.484 , 0.484)	(0.558 , 0.558)
	$A_{fn,avg}$	0.281	0.309	0.377
$<(5,4), \mathbf{G}_p, A>$	E^*_{avg}/L	0.091	0.225	0.304
with MBFD	d^*_{ii}/L	(0.889 , 0.889)	(0.558 , 0.521)	(0.300 , 0.779)
$FAF = 13$dB	$A_{fn,avg}$	0.000	0.037	0.206

Table A2: MSAL efficiency with 2FRS in various localization scenarios

Scenario		$\sigma = 0$	$\sigma = 5$	$\sigma = 10$
		$h = 0.25L$	$h = 0.25L$	$h = 0.25L$
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	13.87	5.15	2.037
	d^*_{ii}/L	(0.853 , 0.705)	(0.374 , 0.374)	(0.300 , 0.447)
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	10.850	5.634	2.488
FP 1, $WAF = 3.1$dB	d^*/L	(0.374 , 0.374)	(0.374 , 0.374)	(0.300 , 0.447)
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	10.357	5.553	2.528
FP 2, $WAF = 3.1$dB	d^*/L	(0.300 , 0.447)	(0.300 , 0.447)	(0.300 , 0.447)
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	12.578	3.447	1.652
$FAF = 13$dB	d^*_{ii}/L	(0.926 , 0.853)	(0.300 , 0.779)	(0.300 , 0.779)
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	10.57145996	3.519244357	1.32113641
with MBFD	d^*_{ii}/L	(0.484 , 0.484)	(0.374 , 0.374)	(0.374 , 0.374)
	$A_{fn,avg}$	0	0	0
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	2.516	0.820	0.317
with MBFD	d^*_{ii}/L	(0.926 , 0.853)	(0.374 , 0.374)	(0.374 , 0.374)
$FAF = 13$dB	$A_{fn,avg}$	0.000	0.020	0.148

Table A3: MSAL accuracy with 3FRS in various localization scenarios

Scenario		$\sigma = 0$	$\sigma = 5$	$\sigma = 10$
		$h = 0.25L$	$h = 0.25L$	$h = 0.25L$
$< (4,3,2), \mathbf{G}_p, A >$	E^*_{avg}/L	0.080	0.179	0.262
	d^*_{ii}/L	(0.521 , 0.521)	(0.484 , 0.447)	(0.595 , 0.521)
$< (4,3,2), \mathbf{G}_p, A, waf >$	E^*_{avg}/L	0.098	0.151	0.211
FP 1, $WAF = 3.1$dB	d^*/L	(0.447 , 0.484)	(0.484 , 0.484)	(0.447 , 0.484)
$< (4,3,2), \mathbf{G}_p, A, waf >$	E^*_{avg}/L	0.088	0.152	0.211
FP 2, $WAF = 3.1$dB	d^*/L	(0.411 , 0.484)	(0.484 , 0.484)	(0.521 , 0.484)
$< (4,3,2), \mathbf{G}_p, A >$	E^*_{avg}/L	0.099	0.264	0.337
$FAF = 13$dB	d^*_{ii}/L	(0.926 , 0.963)	(0.853 , 0.779)	(0.926 , 0.889)
$< (4,3,2), \mathbf{G}_p, A >$	E^*_{avg}/L	0.070	0.170	0.242
with MBFD	d^*_{ii}/L	(0.411 , 0.668)	(0.411 , 0.558)	(0.411 , 0.558)
	$A_{fn,avg}$	0.166	0.343	0.499
$< (4,3,2), \mathbf{G}_p, A >$	E^*_{avg}/L	0.088	0.251	0.313
with MBFD	d^*_{ii}/L	(0.779 , 0.926)	(0.926 , 0.300)	(0.926 , 0.300)
$FAF = 13$dB	$A_{fn,avg}$	0.000	0.235	0.308

Table A4: MSAL efficiency with 3FRS in various localization scenarios

Scenario		$\sigma = 0$	$\sigma = 5$	$\sigma = 10$
		$h = 0.25L$	$h = 0.25L$	$h = 0.25L$
$< (4,3,2), \mathbf{G}_p, A >$	η^*_{avg}	22.752	7.323	2.655
	d^*_{ii}/L	(0.521 , 0.521)	(0.411 , 0.411)	(0.300 , 0.632)
$< (4,3,2), \mathbf{G}_p, A, waf >$	η^*_{avg}	17.055	8.432	3.403
FP 1, $WAF = 3.1$dB	d^*/L	(0.447 , 0.447)	(0.447 , 0.447)	(0.447 , 0.447)
$< (4,3,2), \mathbf{G}_p, A, waf >$	η^*_{avg}	15.452	7.546	3.332
FP 2, $WAF = 3.1$dB	d^*/L	(0.668 , 0.668)	(0.411 , 0.484)	(0.411 , 0.484)
$< (4,3,2), \mathbf{G}_p, A >$	η^*_{avg}	18.738	4.755	2.347
$FAF = 13$dB	d^*_{ii}/L	(0.853 , 1.000)	(0.300 , 1.000)	(0.300 , 1.000)
$< (4,3,2), \mathbf{G}_p, A >$	η^*_{avg}	2.801	0.543	0.157
with MBFD	d^*_{ii}/L	(0.558 , 0.447)	(0.300 , 0.595)	(0.300 , 0.595)
	$A_{fn,avg}$	0.341	0.430	0.529
$< (4,3,2), \mathbf{G}_p, A >$	η^*_{avg}	3.246	0.848	0.313
with MBFD	d^*_{ii}/L	(0.926 , 0.779)	(0.374 , 0.337)	(0.300 , 0.300)
$FAF = 13$dB	$A_{fn,avg}$	0.000	0.038	0.224

Table A5: MSAL+MCA accuracy with 2FRS in various localization scenarios

Scenario		$\sigma = 0$	$\sigma = 5$	$\sigma = 10$
		$h = 0.25L$	$h = 0.25L$	$h = 0.25L$
$<(5,4), \mathbf{G}_p, A>$	E^*_{avg}/L	0.081	0.179	0.279
	d^*_{ii}/L	(0.705 , 0.853)	(0.374 , 0.374)	(0.411 , 0.374)
$<(5,4), \mathbf{G}_p, A>$	E^*_{avg}/L	0.096	0.156	0.231
FP 1, $WAF = 3.1$dB	d^*/L	(0.484 , 0.447)	(0.411 , 0.411)	(0.337 , 0.374)
$<(5,4), \mathbf{G}_p, A>$	E^*_{avg}/L	0.091	0.153	0.222
FP 2, $WAF = 3.1$dB	d^*/L	(0.484 , 0.447)	(0.447 , 0.484)	(0.374 , 0.374)
$<(5,4), \mathbf{G}_p, A>$	E^*_{avg}/L	0.091	0.227	0.314
$FAF = 13$dB	d^*_{ii}/L	(0.889 , 0.889)	(0.521 , 0.521)	(0.558 , 0.632)
$<(5,4), \mathbf{R}, A, sh>$	E^*_{avg}/L	0.087	0.182	
	d^*_{ii}/L	(0.627,0.487)	(0.3,0.3)	
$<(5,4), \mathbf{G}_p, A>$	E^*_{avg}/L	0.077	0.184	0.275
with MBFD	d^*_{ii}/L	(0.484 , 0.484)	(0.447 , 0.447)	(0.374 , 0.374)
	$A_{fn,avg}$	0.281	0.307	0.364
$<(5,4), \mathbf{G}_p, A>$	E^*_{avg}/L	0.091	0.227	0.310
with MBFD	d^*_{ii}/L	(0.889 , 0.889)	(0.521 , 0.521)	(0.521 , 0.521)
$FAF = 13$dB	$A_{fn,avg}$	0.000	0.036	0.179

Table A6: MSAL+MCA efficiency with 2FRS in various localization scenarios

Scenario		$\sigma = 0$	$\sigma = 5$	$\sigma = 10$
		$h = 0.25L$	$h = 0.25L$	$h = 0.25L$
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	13.87	6.13	3.965
	d^*_{ii}/L	(0.853 , 0.705)	(0.411 , 0.411)	(0.447 , 0.484)
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	11.643	7.218	4.793
FP 1, $WAF = 3.1$dB	d^*/L	(0.484 , 0.447)	(0.447 , 0.484)	(0.411 , 0.411)
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	12.391	7.367	5.041
FP 2, $WAF = 3.1$dB	d^*/L	(0.484 , 0.484)	(0.447 , 0.484)	(0.447 , 0.447)
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	12.578	4.885	3.559
$FAF = 13$dB	d^*_{ii}/L	(0.926 , 0.853)	(0.521 , 0.558)	(0.558 , 0.632)
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	10.57145996	4.187882981	2.49337063
with MBFD	d^*_{ii}/L	(0.484 , 0.484)	(0.447 , 0.447)	(0.447 , 0.484)
	$A_{fn,avg}$	0	0	0
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	2.516	0.961	0.594
with MBFD	d^*_{ii}/L	(0.926 , 0.853)	(0.484 , 0.484)	(0.484 , 0.484)
$FAF = 13$dB	$A_{fn,avg}$	0.000	0.034	0.175

Table A7: MSAL+MCA accuracy with 3FRS in various localization scenarios

Scenario		$\sigma = 0$	$\sigma = 5$	$\sigma = 10$
		$h = 0.25L$	$h = 0.25L$	$h = 0.25L$
$< (4,3,2), \mathbf{G}_p, A >$	E^*_{avg}/L	0.080	0.188	0.287
	d^*_{ii}/L	(0.521 , 0.521)	(0.411 , 0.411)	(0.411 , 0.411)
$< (4,3,2), \mathbf{G}_p, A, waf >$	E^*_{avg}/L	0.097	0.157	0.234
FP 1, $WAF = 3.1$dB	d^*/L	(0.447 , 0.484)	(0.411 , 0.374)	(0.411 , 0.374)
$< (4,3,2), \mathbf{G}_p, A, waf >$	E^*_{avg}/L	0.091	0.156	0.225
FP 2, $WAF = 3.1$dB	d^*/L	(0.447 , 0.484)	(0.411 , 0.374)	(0.411 , 0.374)
$< (4,3,2), \mathbf{G}_p, A >$	E^*_{avg}/L	0.099	0.271	0.335
$FAF = 13$dB	d^*_{ii}/L	(0.926 , 0.963)	(0.742 , 0.668)	(0.742 , 0.668)
$< (4,3,2), \mathbf{G}_p, A >$	E^*_{avg}/L	0.070	0.187	0.274
with MBFD	d^*_{ii}/L	(0.411 , 0.668)	(0.337 , 0.447)	(0.411 , 0.411)
	$A_{fn,avg}$	0.166	0.343	0.499
$< (4,3,2), \mathbf{G}_p, A >$	E^*_{avg}/L	0.088	0.257	0.314
with MBFD	d^*_{ii}/L	(0.779 , 0.926)	(0.558 , 0.558)	(0.595 , 0.558)
$FAF = 13$dB	$A_{fn,avg}$	0.000	0.085	0.286

Table A8: MSAL+MCA efficiency with 3FRS in various localization scenarios

Scenario		$\sigma = 0$	$\sigma = 5$	$\sigma = 10$
		$h = 0.25L$	$h = 0.25L$	$h = 0.25L$
$< (4,3,2), \mathbf{G}_p, A >$	η^*_{avg}	22.752	9.453	6.219
	d^*_{ii}/L	(0.521 , 0.521)	(0.447 , 0.447)	(0.447 , 0.484)
$< (4,3,2), \mathbf{G}_p, A, waf >$	η^*_{avg}	18.895	11.348	7.594
FP 1, $WAF = 3.1$dB	d^*/L	(0.484 , 0.447)	(0.374 , 0.411)	(0.374 , 0.411)
$< (4,3,2), \mathbf{G}_p, A, waf >$	η^*_{avg}	19.701	11.541	7.942
FP 2, $WAF = 3.1$dB	d^*/L	(0.484 , 0.447)	(0.374 , 0.411)	(0.374 , 0.411)
$< (4,3,2), \mathbf{G}_p, A >$	η^*_{avg}	18.738	6.795	5.435
$FAF = 13$dB	d^*_{ii}/L	(0.853 , 1.000)	(0.742 , 0.668)	(0.742 , 0.668)
$< (4,3,2), \mathbf{G}_p, A >$	η^*_{avg}	2.801	0.879	0.503
with MBFD	d^*_{ii}/L	(0.558 , 0.447)	(0.411 , 0.374)	(0.374 , 0.337)
	$A_{fn,avg}$	0.341	0.423	0.521
$< (4,3,2), \mathbf{G}_p, A >$	η^*_{avg}	3.246	1.042	0.657
with MBFD	d^*_{ii}/L	(0.926 , 0.779)	(0.484 , 0.521)	(0.447 , 0.632)
$FAF = 13$dB	$A_{fn,avg}$	0.000	0.076	0.285

Table A9: MCA-iso accuracy with 2FRS in various localization scenarios

Scenario		$\sigma = 0$	$\sigma = 5$	$\sigma = 10$
		$h = 0.25L$	$h = 0.25L$	$h = 0.25L$
$< (5,4), \mathbf{G}_p, A >$	E^*_{avg}/L	0.090	0.189	0.284
	d^*_{ii}/L	(0.484 , 0.484)	(0.484 , 0.447)	(0.521 , 0.411)
$< (5,4), \mathbf{G}_p, A >$	E^*_{avg}/L	0.100	0.161	0.238
FP 1, $WAF = 3.1$dB	d^*_{ii}/L	(0.447 , 0.447)	(0.447 , 0.521)	(0.447 , 0.521)
$< (5,4), \mathbf{G}_p, A >$	E^*_{avg}/L	0.092	0.156	0.226
FP 2, $WAF = 3.1$dB	d^*_{ii}/L	(0.484 , 0.411)	(0.447 , 0.595)	(0.411 , 0.521)
$< (5,4), \mathbf{G}_p, A >$	E^*_{avg}/L	0.123	0.233	0.315
	d^*_{ii}/L	(0.816 , 0.447)	(0.558 , 0.558)	(0.558 , 0.779)
$< (5,4), \mathbf{R}, A, sh >$	E^*_{avg}/L	0.141	0.189	
$FAF = 13$dB	d^*_{ii}/L	(0.300 0.300)	(0.300 0.300)	
$< (5,4), \mathbf{G}_p, A >$	E^*_{avg}/L	0.083	0.190	0.281
with MBFD	d^*_{ii}/L	(0.484 , 0.484)	(0.484 , 0.484)	(0.521 , 0.374)
	$A_{fn,avg}$	0.281	0.309	0.374
$< (5,4), \mathbf{G}_p, A >$	E^*_{avg}/L	0.122	0.233	0.311
with MBFD	d^*_{ii}/L	(0.300 , 0.816)	(0.558 , 0.595)	(0.705 , 0.411)
$FAF = 13$dB	$A_{fn,avg}$	0.220	0.038	0.183

Table A10: MCA-iso efficiency with 2FRS in various localization scenarios

Scenario		$\sigma = 0$	$\sigma = 5$	$\sigma = 10$
		$h = 0.25L$	$h = 0.25L$	$h = 0.25L$
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	12.494	5.931	3.929
	d^*_{ii}/L	(0.484 , 0.447)	(0.447 , 0.484)	(0.411 , 0.521)
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	11.226	6.989	4.709
FP 1, $WAF = 3.1$dB	d^*_{ii}/L	(0.521 , 0.447)	(0.447 , 0.521)	(0.447 , 0.521)
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	12.370	7.220	4.968
FP 2, $WAF = 3.1$dB	d^*_{ii}/L	(0.484 , 0.411)	(0.447 , 0.595)	(0.411 , 0.521)
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	10.430	4.789	3.560
$FAF = 13$dB	d^*_{ii}/L	(0.816 , 0.447)	(0.558 , 0.595)	(0.558 , 0.779)
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	9.710030845	4.06669776	2.46991029
with MBFD	d^*_{ii}/L	(0.484 , 0.484)	(0.484 , 0.521)	(0.521 , 0.521)
	$A_{fn,avg}$	0	0	0
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	2.034	0.926	0.586
with MBFD	d^*_{ii}/L	(0.816 , 0.447)	(0.595 , 0.558)	(0.521 , 0.632)
$FAF = 13$dB	$A_{fn,avg}$	0.012	0.038	0.181

Table A11: MCA-iso accuracy with 3FRS in various localization scenarios

Scenario		$\sigma = 0$	$\sigma = 5$	$\sigma = 10$
		$h = 0.25L$	$h = 0.25L$	$h = 0.25L$
$< (4,3,2), \mathbf{G}_p, A >$	E^*_{avg}/L	0.094	0.195	0.290
	d^*_{ii}/L	(0.484 , 0.484)	(0.521 , 0.484)	(0.484 , 0.558)
$< (4,3,2), \mathbf{G}_p, A, waf >$	E^*_{avg}/L	0.100	0.164	0.242
FP 1, $WAF = 3.1$dB	d^*_{ii}/L	(0.447 , 0.558)	(0.558 , 0.521)	(0.484 , 0.484)
$< (4,3,2), \mathbf{G}_p, A, waf >$	E^*_{avg}/L	0.094	0.160	0.230
FP 2, $WAF = 3.1$dB	d^*_{ii}/L	(0.447 , 0.484)	(0.558 , 0.484)	(0.558 , 0.447)
$< (4,3,2), \mathbf{G}_p, A >$	E^*_{avg}/L	0.139	0.270	0.333
$FAF = 13$dB	d^*_{ii}/L	(0.816 , 0.816)	(0.705 , 0.742)	(0.816 , 0.816)
$< (4,3,2), \mathbf{G}_p, A >$	E^*_{avg}/L	0.110	0.189	0.273
with MBFD	d^*_{ii}/L	(0.484 , 0.484)	(0.411 , 0.447)	(0.411 , 0.411)
	$A_{fn,avg}$	0.166	0.343	0.499
$< (4,3,2), \mathbf{G}_p, A >$	E^*_{avg}/L	0.165	0.259	0.314
with MBFD	d^*_{ii}/L	(0.374 , 0.816)	(0.558 , 0.558)	(0.963 , 0.300)
$FAF = 13$dB	$A_{fn,avg}$	0.105	0.085	0.307

Table A12: MCA-iso efficiency with 3FRS in various localization scenarios

Scenario		$\sigma = 0$	$\sigma = 5$	$\sigma = 10$
		$h = 0.25L$	$h = 0.25L$	$h = 0.25L$
$< (4,3,2), \mathbf{G}_p, A >$	η^*_{avg}	19.270	9.245	6.211
	d^*_{ii}/L	(0.484 , 0.484)	(0.521 , 0.484)	(0.558 , 0.595)
$< (4,3,2), \mathbf{G}_p, A, waf >$	η^*_{avg}	17.896	10.986	7.429
FP 1, $WAF = 3.1$dB	d^*_{ii}/L	(0.558 , 0.447)	(0.521 , 0.558)	(0.484 , 0.484)
$< (4,3,2), \mathbf{G}_p, A, waf >$	η^*_{avg}	19.045	11.277	7.841
FP 2, $WAF = 3.1$dB	d^*_{ii}/L	(0.484 , 0.447)	(0.484 , 0.558)	(0.484 , 0.632)
$< (4,3,2), \mathbf{G}_p, A >$	η^*_{avg}	14.614	6.785	5.454
$FAF = 13$dB	d^*_{ii}/L	(0.816 , 0.447)	(0.742 , 0.779)	(0.816 , 0.926)
$< (4,3,2), \mathbf{G}_p, A >$	η^*_{avg}	1.810	0.864	0.497
with MBFD	d^*_{ii}/L	(0.484 , 0.484)	(0.447 , 0.374)	(0.374 , 0.337)
	$A_{fn,avg}$	0.341	0.431	0.521
$< (4,3,2), \mathbf{G}_p, A >$	η^*_{avg}	1.813	1.031	0.654
with MBFD	d^*_{ii}/L	(0.779 , 0.484)	(0.595 , 0.521)	(0.558 , 0.595)
$FAF = 13$dB	$A_{fn,avg}$	0.026	0.088	0.286

Table A13: MSAL+MCA-dpl accuracy with 2FRS in various localization scenarios

Scenario		$\sigma = 0$	$\sigma = 5$	$\sigma = 10$
		$h = 0.25L$	$h = 0.25L$	$h = 0.25L$
$< (5,4), \mathbf{G}_p, A >$	E^*_{avg}/L	0.081	0.218	0.319
	d^*_{ii}/L	(0.889 , 0.742)	(0.558 , 0.558)	(0.521 , 0.558)
$< (5,4), \mathbf{G}_p, A >$	E^*_{avg}/L	0.115	0.193	0.251
FP 1, $WAF = 3.1$dB	d^*_{ii}/L	(0.558 , 0.595)	(0.300 , 0.374)	(0.300 , 0.300)
$< (5,4), \mathbf{G}_p, A >$	E^*_{avg}/L	0.114	0.191	0.247
FP 2, $WAF = 3.1$dB	d^*_{ii}/L	(0.668 , 0.595)	(0.484 , 0.447)	(0.300 , 0.300)
$< (5,4), \mathbf{G}_p, A >$	E^*_{avg}/L	0.137	0.230	0.316
$FAF = 13$dB	d^*_{ii}/L	(0.742 , 0.779)	(0.521 , 0.521)	(0.521 , 0.558)
$< (5,4), \mathbf{G}_p, A >$	E^*_{avg}/L	0.082	0.217	0.314
with MBFD	d^*_{ii}/L	(0.742 , 0.889)	(0.558 , 0.558)	(0.558 , 0.521)
	$A_{fn,avg}$	0.233	0.107	0.069
$< (5,4), \mathbf{G}_p, A >$	E^*_{avg}/L	0.137	0.230	0.316
with MBFD	d^*_{ii}/L	(0.742 , 0.779)	(0.521 , 0.521)	(0.521 , 0.521)
$FAF = 13$dB	$A_{fn,avg}$	0	0	0

Table A14: MSAL+MCA-dpl efficiency with 2FRS in various localization scenarios

Scenario		$\sigma = 0$	$\sigma = 5$	$\sigma = 10$
		$h = 0.25L$	$h = 0.25L$	$h = 0.25L$
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	13.874	5.125	3.471
	d^*_{ii}/L	(0.889 , 0.742)	(0.558 , 0.558)	(0.521 , 0.595)
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	9.834	5.744	4.057
FP 1, $WAF = 3.1$dB	d^*/L	(0.595 , 0.558)	(0.558 , 0.558)	(0.484 , 0.484)
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	9.838	5.792	4.214
FP 2, $WAF = 3.1$dB	d^*/L	(0.595 , 0.668)	(0.447 , 0.484)	(0.484 , 0.484)
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	8.213	4.798	3.460
$FAF = 13$dB	d^*_{ii}/L	(0.779 , 0.742)	(0.521 , 0.558)	(0.668 , 0.595)
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	2.126	0.985	0.699
with MBFD	d^*_{ii}/L	(0.926 , 0.668)	(0.484 , 0.484)	(0.447 , 0.484)
	$A_{fn,avg}$	0.233	0.107	0.069
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	1.643	0.986	0.701
with MBFD	d^*_{ii}/L	(0.779 , 0.742)	(0.484 , 0.484)	(0.484 , 0.484)
$FAF = 13$dB	$A_{fn,avg}$	0	0	0

Table A15: MSAL+MCA-dpl accuracy with 3FRS in various localization scenarios

Scenario		$\sigma = 0$	$\sigma = 5$	$\sigma = 10$
		$h = 0.25L$	$h = 0.25L$	$h = 0.25L$
$< (4,3,2), \mathbf{G}_p, A >$	η^*_{avg}	0.083	0.241	0.350
	d^*_{ii}/L	(0.779 , 0.779)	(0.632 , 0.632)	(0.816 , 0.742)
$< (4,3,2), \mathbf{G}_p, A, waf >$	η^*_{avg}	0.130	0.213	0.286
FP 1, $WAF = 3.1$dB	d^*/L	(0.595 , 0.632)	(0.300 , 0.300)	(0.300 , 0.300)
$< (4,3,2), \mathbf{G}_p, A, waf >$	η^*_{avg}	0.128	0.217	0.283
FP 2, $WAF = 3.1$dB	d^*/L	(0.668 , 0.668)	(0.300 , 0.300)	(0.300 , 0.300)
$< (4,3,2), \mathbf{G}_p, A >$	η^*_{avg}	0.230	0.297	0.344
$FAF = 13$dB	d^*_{ii}/L	(0.705 , 0.742)	(0.705 , 0.705)	(0.705 , 0.742)
$< (4,3,2), \mathbf{G}_p, A >$	η^*_{avg}	0.076	0.223	0.287
with MBFD	d^*_{ii}/L	(0.742 , 0.779)	(0.558 , 0.595)	(0.668 , 0.300)
	$A_{fn,avg}$	0	0.23725	0.46108
$< (4,3,2), \mathbf{G}_p, A >$	η^*_{avg}	0.230	0.278	0.334
with MBFD	d^*_{ii}/L	(0.705 , 0.742)	(0.558 , 0.558)	(0.558 , 0.558)
$FAF = 13$dB	$A_{fn,avg}$	0.000	0.000	0.000

Table A16: MSAL+MCA-dpl efficiency with 3FRS in various localization scenarios

Scenario		$\sigma = 0$	$\sigma = 5$	$\sigma = 10$
		$h = 0.25L$	$h = 0.25L$	$h = 0.25L$
$< (4,3,2), \mathbf{G}_p, A >$	E^*_{avg}/L	22.274	7.519	5.201
	d^*_{ii}/L	(0.816 , 0.779)	(0.632 , 0.632)	(0.816 , 0.742)
$< (4,3,2), \mathbf{G}_p, A, waf >$	E^*_{avg}/L	14.480	8.090	5.644
FP 1, $WAF = 3.1$dB	d^*/L	(0.668 , 0.595)	(0.595 , 0.595)	(0.595 , 0.558)
$< (4,3,2), \mathbf{G}_p, A, waf >$	E^*_{avg}/L	14.116	8.088	5.832
FP 2, $WAF = 3.1$dB	d^*/L	(0.668 , 0.668)	(0.595 , 0.595)	(0.595 , 0.595)
$< (4,3,2), \mathbf{G}_p, A >$	E^*_{avg}/L	9.670	6.234	5.010
$FAF = 13$dB	d^*_{ii}/L	(0.742 , 0.705)	(0.705 , 0.742)	(0.742 , 0.779)
$< (4,3,2), \mathbf{G}_p, A >$	E^*_{avg}/L	2.813	0.882	0.511
with MBFD	d^*_{ii}/L	(0.742 , 0.595)	(0.447 , 0.447)	(0.447 , 0.411)
	$A_{fn,avg}$	0.268	0.237	0.462
$< (4,3,2), \mathbf{G}_p, A >$	E^*_{avg}/L	1.464	1.056	0.828
with MBFD	d^*_{ii}/L	(0.742 , 0.411)	(0.484 , 0.521)	(0.484 , 0.558)
$FAF = 13$dB	$A_{fn,avg}$	0.000	0.000	0.000

Table A17: MCA-dpl accuracy with 2FRS in various localization scenarios

Scenario		$\sigma = 0$	$\sigma = 5$	$\sigma = 10$
		$h = 0.25L$	$h = 0.25L$	$h = 0.25L$
$< (5,4), \mathbf{G}_p, A >$	E^*_{avg}/L	0.117	0.221	0.320
	d^*_{ii}/L	(0.558 , 0.558)	(0.595 , 0.558)	(0.558 , 0.668)
$< (5,4), \mathbf{G}_p, A >$	E^*_{avg}/L	0.127	0.203	0.286
FP 1, $WAF = 3.1$dB	d^*_{ii}/L	(0.558 , 0.595)	(0.595 , 0.595)	(0.558 , 0.484)
$< (5,4), \mathbf{G}_p, A >$	E^*_{avg}/L	0.121	0.202	0.275
FP 2, $WAF = 3.1$dB	d^*_{ii}/L	(0.668 , 0.595)	(0.632 , 0.558)	(0.558 , 0.484)
$< (5,4), \mathbf{G}_p, A >$	E^*_{avg}/L	0.144	0.239	0.321
$FAF = 13$dB	d^*_{ii}/L	(0.595 , 0.595)	(0.558 , 0.558)	(0.521 , 0.779)
$< (5,4), \mathbf{G}_p, A >$	E^*_{avg}/L	0.124	0.222	0.315
with MBFD	d^*_{ii}/L	(0.558 , 0.558)	(0.558 , 0.595)	(0.632 , 0.521)
	$A_{fn,avg}$	0.132	0.114	0.078
$< (5,4), \mathbf{G}_p, A >$	E^*_{avg}/L	0.144	0.239	0.321
with MBFD	d^*_{ii}/L	(0.595 , 0.595)	(0.558 , 0.558)	(0.632 , 0.742)
$FAF = 13$dB	$A_{fn,avg}$	0	0	0

Table A18: MCA-dpl efficiency with 2FRS in various localization scenarios

Scenario		$\sigma = 0$	$\sigma = 5$	$\sigma = 10$
		$h = 0.25L$	$h = 0.25L$	$h = 0.25L$
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	9.588	5.082	3.501
	d^*_{ii}/L	(0.558 , 0.558)	(0.595 , 0.558)	(0.595 , 0.668)
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	8.856	5.518	3.885
FP 1, $WAF = 3.1$dB	d^*_{ii}/L	(0.595 , 0.558)	(0.595 , 0.595)	(0.558 , 0.632)
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	9.267	5.561	4.052
FP 2, $WAF = 3.1$dB	d^*_{ii}/L	(0.595 , 0.668)	(0.558 , 0.632)	(0.484 , 0.668)
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	7.821	4.664	3.477
$FAF = 13$dB	d^*_{ii}/L	(0.595 , 0.595)	(0.558 , 0.558)	(0.779 , 0.742)
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	1.735	0.918	0.671
with MBFD	d^*_{ii}/L	(0.595 , 0.374)	(0.484 , 0.558)	(0.484 , 0.668)
	$A_{fn,avg}$	0.132	0.114	0.078
$< (5,4), \mathbf{G}_p, A >$	η^*_{avg}	1.564	0.933	0.694
with MBFD	d^*_{ii}/L	(0.595 , 0.595)	(0.558 , 0.558)	(0.853 , 0.705)
$FAF = 13$dB	$A_{fn,avg}$	0	0	0

Table A19: MCA-dpl accuracy with 3FRS in various localization scenarios

Scenario		$\sigma = 0$	$\sigma = 5$	$\sigma = 10$
		$h = 0.25L$	$h = 0.25L$	$h = 0.25L$
$< (4,3,2), \mathbf{G}_p, A >$	E^*_{avg}/L	0.126	0.239	0.343
	d^*_{ii}/L	(0.742 , 0.705)	(0.632 , 0.632)	(0.705 , 0.816)
$< (4,3,2), \mathbf{G}_p, A, waf >$	E^*_{avg}/L	0.139	0.229	0.321
FP 1, $WAF = 3.1$dB	d^*_{ii}/L	(0.558 , 0.668)	(0.632 , 0.668)	(0.705 , 0.742)
$< (4,3,2), \mathbf{G}_p, A, waf >$	E^*_{avg}/L	0.141	0.230	0.310
FP 2, $WAF = 3.1$dB	d^*_{ii}/L	(0.705 , 0.632)	(0.705 , 0.668)	(0.705 , 0.779)
$< (4,3,2), \mathbf{G}_p, A >$	E^*_{avg}/L	0.261	0.330	0.371
$FAF = 13$dB	d^*_{ii}/L	(0.779 , 0.816)	(0.853 , 0.816)	(0.926 , 1.000)
$< (4,3,2), \mathbf{G}_p, A >$	E^*_{avg}/L	0.150	0.224	0.290
with MBFD	d^*_{ii}/L	(0.300 , 0.742)	(0.300 , 0.668)	(0.668 , 0.300)
	$A_{fn,avg}$	0	0.23725	0.46108
$< (4,3,2), \mathbf{G}_p, A >$	E^*_{avg}/L	0.234	0.283	0.340
with MBFD	d^*_{ii}/L	(0.632 , 0.632)	(0.558 , 0.558)	(0.558 , 0.595)
$FAF = 13$dB	$A_{fn,avg}$	0.000	0.000	0.000

Table A20: MCA-dpl efficiency with 3FRS in various localization scenarios

Scenario		$\sigma = 0$	$\sigma = 5$	$\sigma = 10$
		$h = 0.25L$	$h = 0.25L$	$h = 0.25L$
$< (4,3,2), \mathbf{G}_p, A >$	η^*_{avg}	14.553	7.582	5.292
	d^*_{ii}/L	(0.742 , 0.705)	(0.632 , 0.668)	(0.779 , 0.779)
$< (4,3,2), \mathbf{G}_p, A, waf >$	η^*_{avg}	14.008	8.011	5.693
FP 1, $WAF = 3.1$dB	d^*_{ii}/L	(0.668 , 0.558)	(0.668 , 0.668)	(0.742 , 0.742)
$< (4,3,2), \mathbf{G}_p, A, waf >$	η^*_{avg}	13.035	8.044	5.926
FP 2, $WAF = 3.1$dB	d^*_{ii}/L	(0.632 , 0.705)	(0.668 , 0.705)	(0.779 , 0.705)
$< (4,3,2), \mathbf{G}_p, A >$	η^*_{avg}	8.082	5.553	4.782
$FAF = 13$dB	d^*_{ii}/L	(0.853 , 0.779)	(0.889 , 0.926)	(1.000 , 1.000)
$< (4,3,2), \mathbf{G}_p, A >$	η^*_{avg}	1.570	0.830	0.500
with MBFD	d^*_{ii}/L	(0.558 , 0.374)	(0.484 , 0.484)	(0.521 , 0.521)
	$A_{fn,avg}$	0.169	0.240	0.472
$< (4,3,2), \mathbf{G}_p, A >$	η^*_{avg}	1.415	1.022	0.837
with MBFD	d^*_{ii}/L	(0.632 , 0.447)	(0.595 , 0.632)	(0.779 , 0.926)
$FAF = 13$dB	$A_{fn,avg}$	0.000	0.000	0.000

A.2 MULTI-FLOOR LOCALIZATION & FLOOR DETERMINATION WITH 2FRS

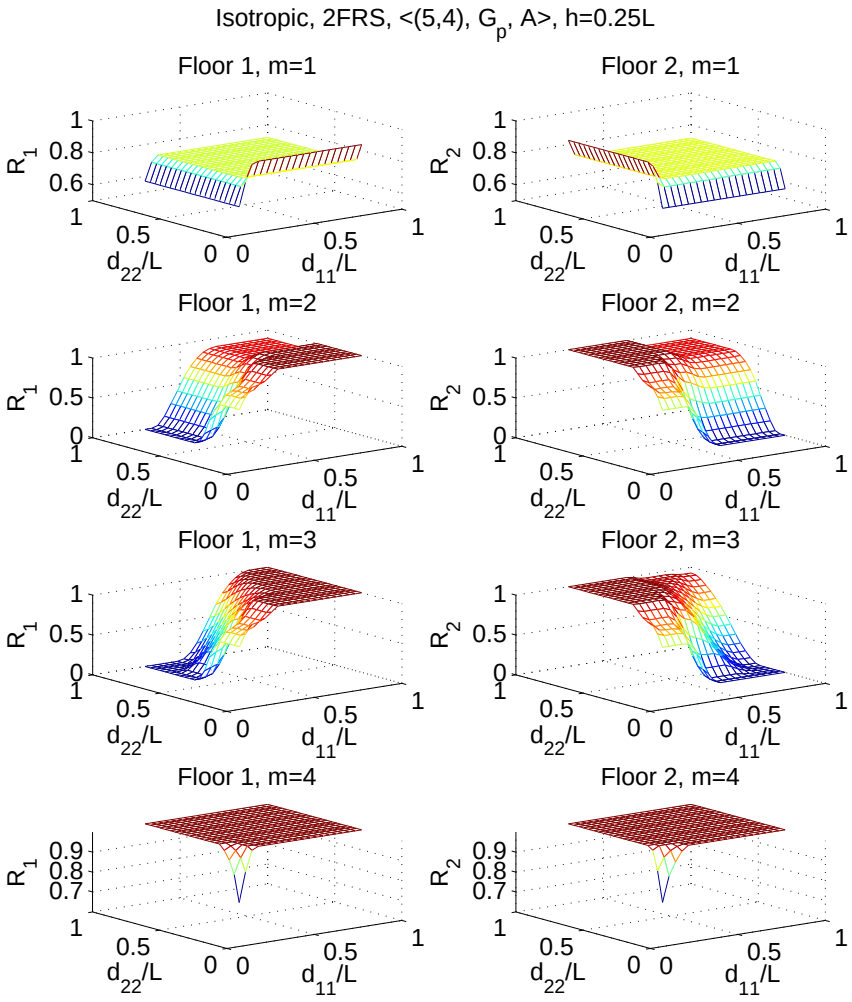

Figure A6: Availability on floors 1 and 2 with MSAL-iso, $\sigma = 0$, $h = 0.25L$

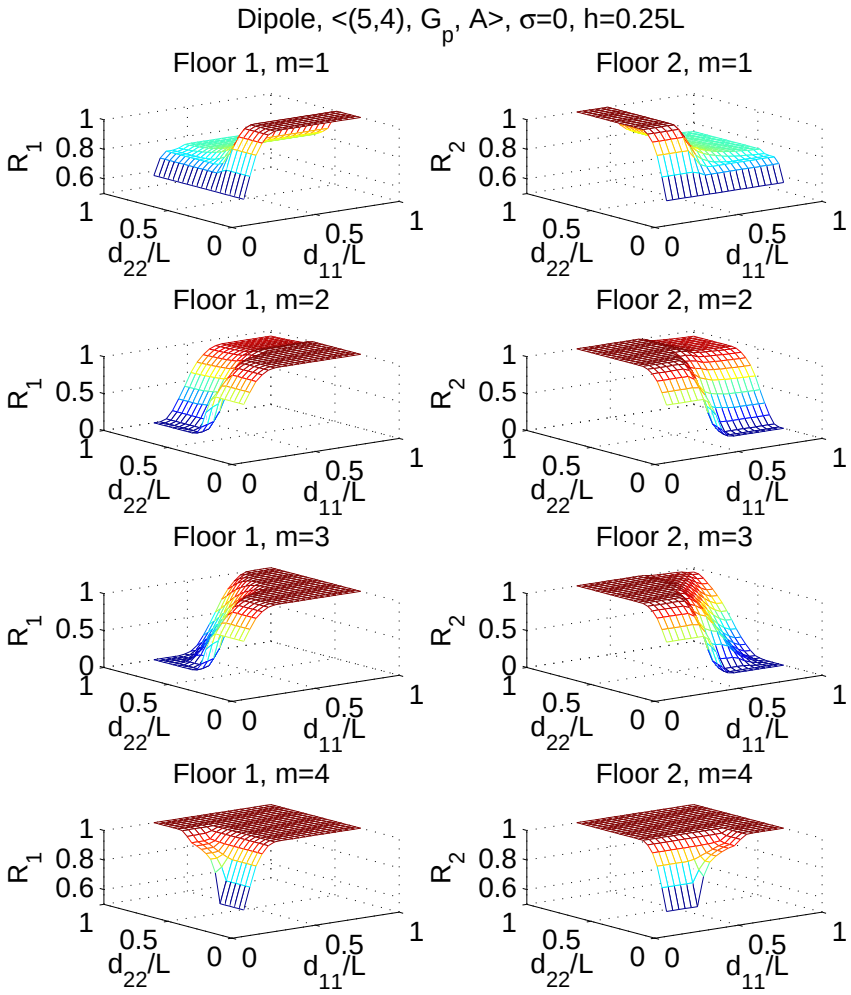

Figure A7: Availability on floors 1 and 2 with MSAL-dpl, $\sigma = 0$, $h = 0.25L$

A.3 MULTI-FLOOR LOCALIZATION & FLOOR DETERMINATION WITH 3FRS

In this section, we provide a detailed performance analysis on FD performance of a 3 floor repetition localization scenario (3FRS) under various channel conditions and with different antenna types. Later in Section A.3.1, the performance of multi-floor subarea localization and multi-floor centroid algorithm with 3FRS are shown when integrated with FD algorithm.

While evaluating the floor determination for 3 floor scheme, we also want to show the performance under the special case where the highest RSS readings in multiple floors are compared to each other to determine the correct floor. In the Section 4.4.1.1, we listed 2 constraints that must be satisfied for no false negatives. 3FRS scheme introduced in previous sections use a total of 9 MoSs in 3 floors, $< (4,3,2), \mathbf{G}_p, A >$. In this section, we modify 3FRS scheme so that these two constraints are satisfied assuming isotropic antennas are used on MoSs. Modified 3FRS $< (5,3,3), \mathbf{G}_p, A >$ localization scenario shown in Figure A8 has one MoS at the center of each floor and uses 3.6 MoSs on the average per floor instead of 3 MoSs on the average in scenario $< (4,3,2), \mathbf{G}_p, A >$. The first constraint is satisfied if $d_{ii} \geq L/2$ for floors 1 and 4, and if $d_{ii} \geq \sqrt{5}L/4$ for floors 2,3 and 5. Second constraint is satisfied when monitoring ranges are adjusted so that the conditions are valid under Theorem 1 and 2 with given floor height.

We show the percentage of false negatives with isotropic antennas when $h = 0.5L$ with 3FRS scenario in Figure A9. In this case, it is obvious that when $d_{44} \geq L/2$ and $d_{22} == d_{33} \geq \sqrt{5}L/4$, the first constraint is satisfied. When $m = 1$ at $h = 0.5L$, second constraint is also satisfied, and probability of false negatives is 0. As m increases, the range that can provide 0 false negatives probability narrows down because multiple RSS values from each floor are combined together and compared against each other for FD. In 3FRS scenario, for every MN, m RSS measurements are collected from 3 different floors and sum of these values are compared to each other to find the correct floor among the three floors. From Theorem 2, case 1, and Figure **??**, we can observe that with $h = 0.35L$, 0 false negatives can be achieved when $d < 0.86$, assuming d satisfies the first condition. Therefore, in Figure A10, we show the A_{fn} on three floors using dipole antennas when $h = 0.35L$. We remind

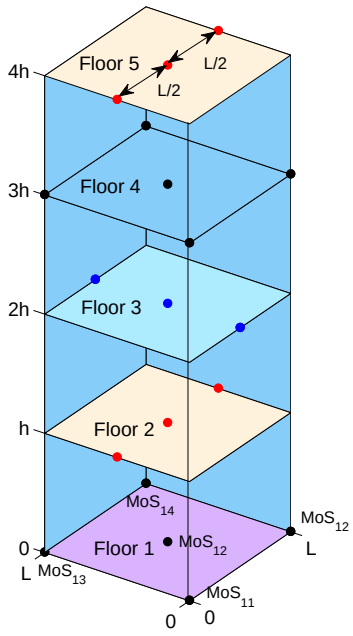

Figure A8: 3FRS, $< (5, 3, 3), \mathbf{G}_p, A >$

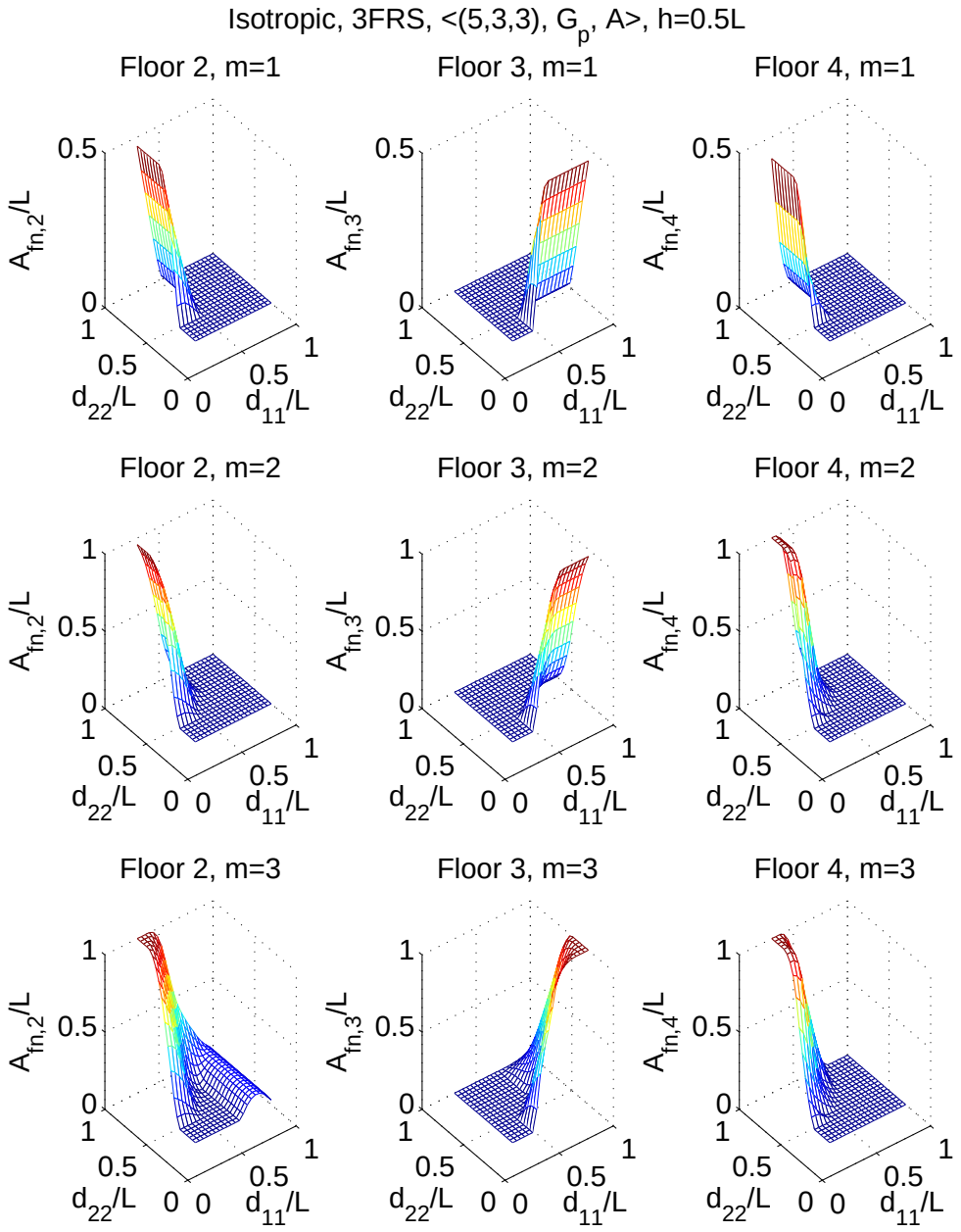

Figure A9: Percentage of false negatives with $< (5,3,3), \mathbf{G}_p, A, iso >$, $\sigma = 0$, $h = 0.5L$

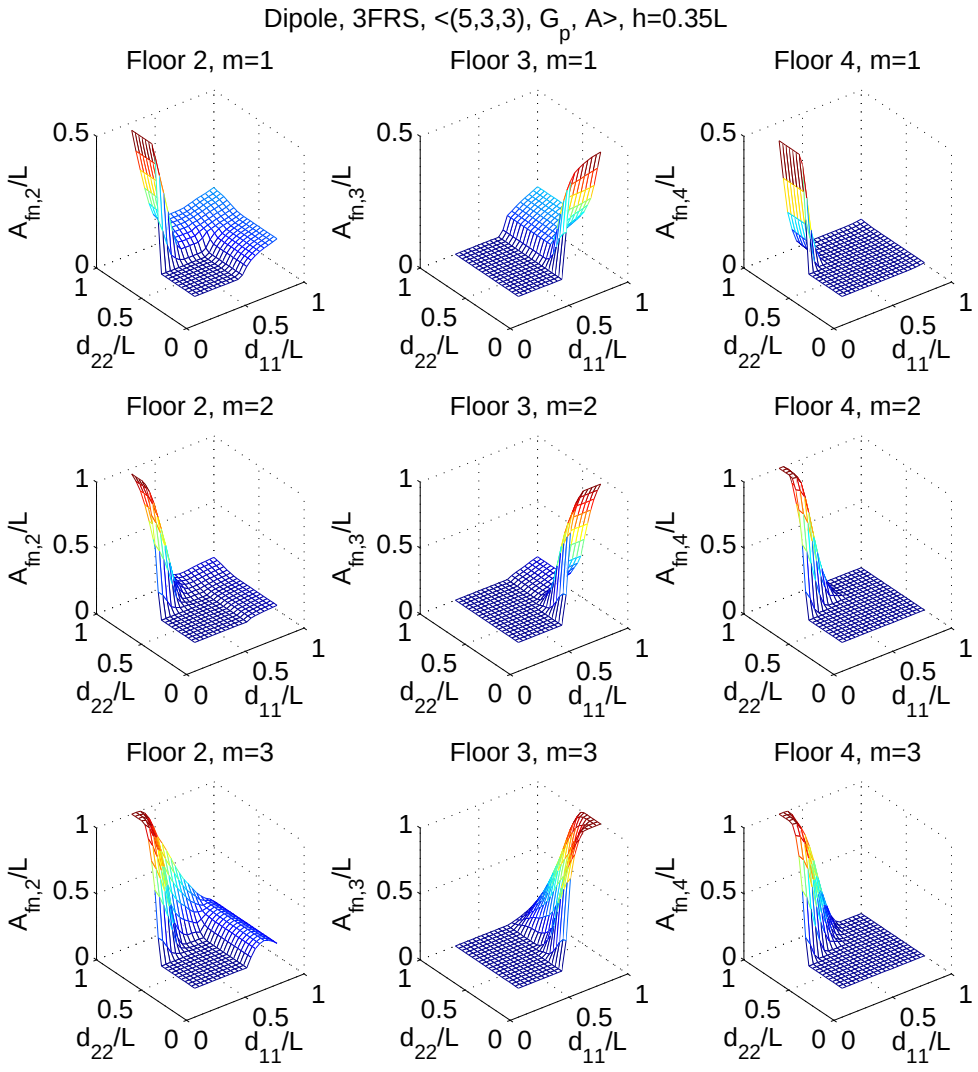

Figure A10: Percentage of false negatives with $< (5,3,3), \mathbf{G}_p, A, iso >$, $\sigma = 0$, $h = 0.35L$

that for adjacent monitoring ranges to be non-zero with dipole antennas, $d_{ii} > 2h$ must be satisfied. As case 1 of theorem 1 is applicable in Floor 4, monitoring ranges that satisfies both condition 1 and condition in Theorem 1-case1 provides 0 false negatives on Floor 4.

We determine the smallest possible fraction of false negatives A^*_{fn} over the possible monitoring ranges in each floor. Figures A11 and A12 show the *minimum* average percentage of false negatives, A^*_{fn} that can be achieved for a given decision depth m, floor height, h, and the standard deviation of the shadowing, σ with 3FRS scenario, where:

$$A^*_{fn} = min_{d_{ii}, i=2...4} \left(\frac{1}{3} \sum_{i=1}^{F} A_{fn,i} \right). \tag{.1}$$

In these figures, we ensure that the 1st constraint is satisfied, i.e., under ideal conditions, full coverage on every floor is provided by adjusting the monitoring ranges. In general, we observe increasing number of false negatives for increasing floor height and shadowing effect. The effect of floor height arises from the adjacent monitoring range depending on floor height as discussed in Theorems 1 and 2. Shadowing effect disrupts the adjacent monitoring range, therefore, false negative probability gets higher. With dipole antennas, we observe $A^*_{fn} = 0$ when $h \leq 0.3L$ instead of $0.35L$ as discussed previously. the reason is that at $0.3L$, the first constraint is satisfied, but the monitoring range $d < 2h$, therefore, adjacent monitoring ranges do not exist, then, in floor determination each MN is detected by MoSs on the floor they reside in. Multiple RSS readings cannot be collected from multiple floors.

To get a more detailed look, Figure A13 shows minimum false negatives probability with respect to floor height under ideal conditions. According to analysis, probability of false negatives is 0 at decision depth $m = 1$ when $h \geq 0.5L$ and $\sigma = 0$. As also shown with the conditions, at $h = 0.5L$, decision depth $m = 1$ can provide zero false negatives under ideal conditions. However, using a larger decision depth can provide a lower A_{fn} value for smaller floor height. We also observe that floor height has a sharper effect on false negatives with dipole antennas. This is due to smaller vertical range of dipole antennas. The reason that we observe 0 false negatives with floor height larger than $0.3L$ is explained before. In contrast to the effect of decision depth with isotropic antennas, with dipole antennas higher decision depth values result in higher percentage of false negatives.

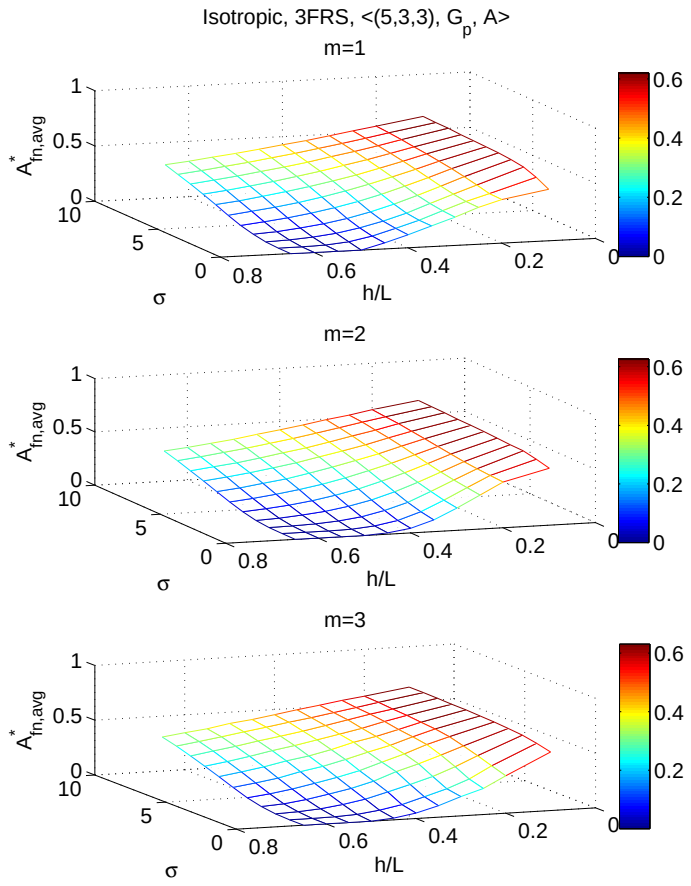

Figure A11: Minimum false negative probability with $< (5,3,3), \mathbf{G}_p, A, iso >$

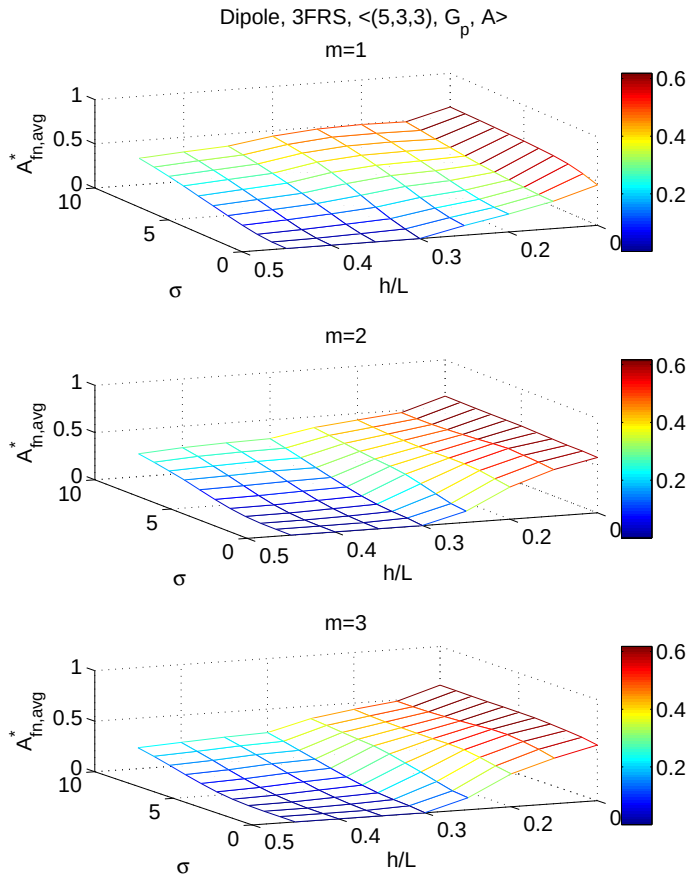

Figure A12: Minimum false negative probability with $< (5,3,3), \mathbf{G}_p, A, dpl >$

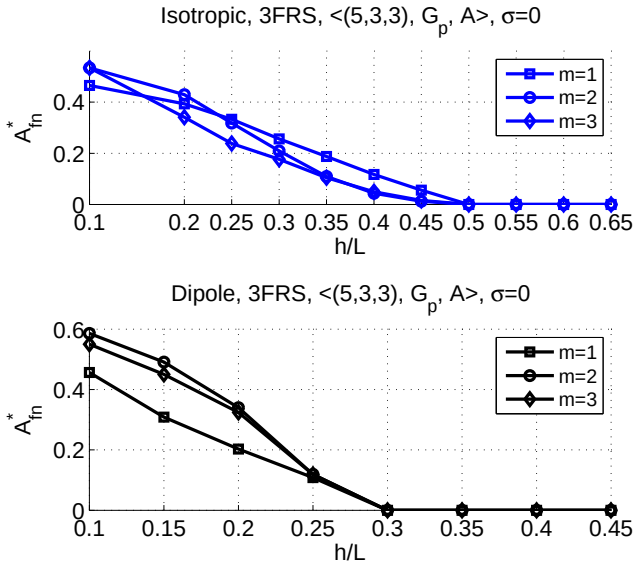

Figure A13: Percentage of false negatives with $< (5,3,3), \mathbf{G}_p, A, iso/dpl >$, $\sigma = 0$

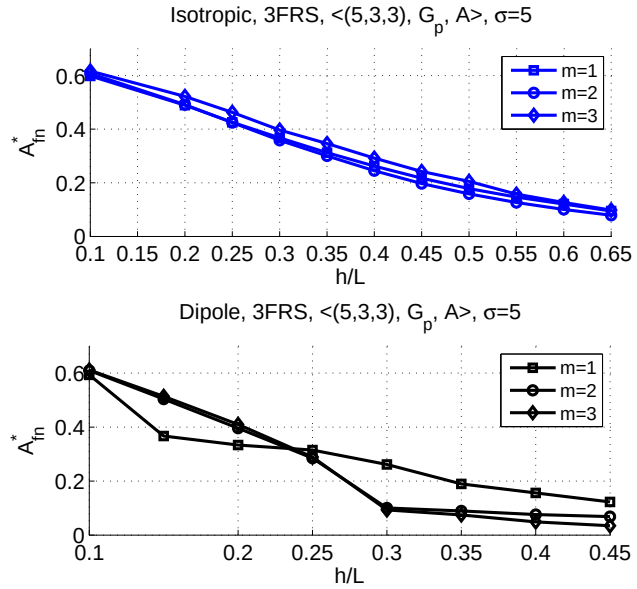

Figure A14: Probability of false negatives with $< (5,3,3), \mathbf{G}_p, A, iso/dpl >$, $\sigma = 5$

Figure A15 shows the minimum false negatives with respect to σ when $h = 0.25L$. With isotropic antennas, we observe A^*_{fn} with decision depth $m = 1$ is higher than with $m = 3$ when $\sigma \leq 2$. Under higher shadowing effect, $m = 1$ gives the smallest percentage of false negatives. With $m = 1$, the reason of smaller A_{fn} is that as σ is increasing monitoring ranges are adjusted so that only one floor's MoSs monitor adjacent floors, the other floor's MoSs only monitor their current floor. We note that A^*_{fn} with dipole antennas is much smaller than A^*_{fn} with isotropic antennas due to smaller vertical range on dipole antennas when compared to isotropic antennas. As σ is getting larger, we observe A^*_{fn} continues to increase, and best A^*_{fn} is given by decision depth $m = 2$.

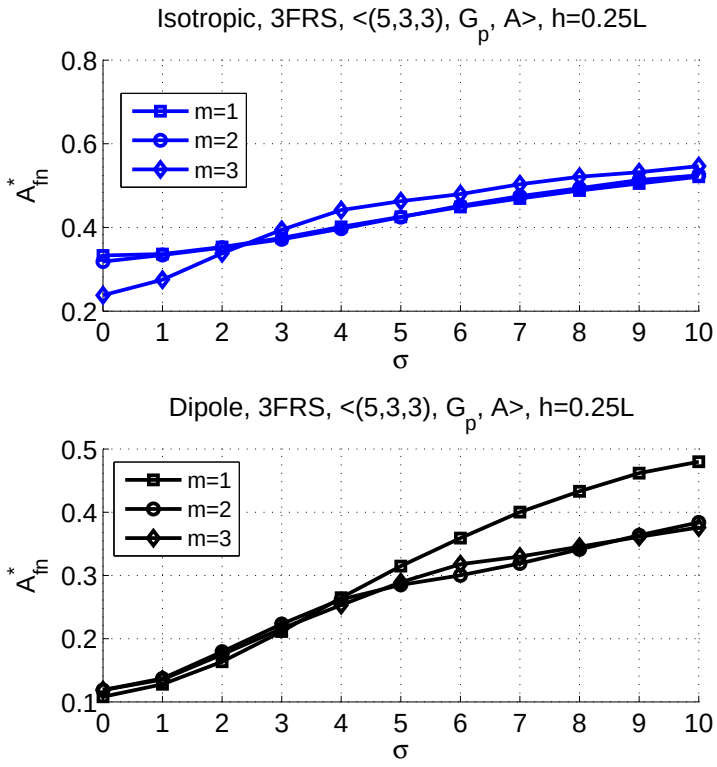

Figure A15: Minimum false negative probability with $< (5, 3, 3), \mathbf{G}_p, A, iso/dpl >$, $h = 0.25L$

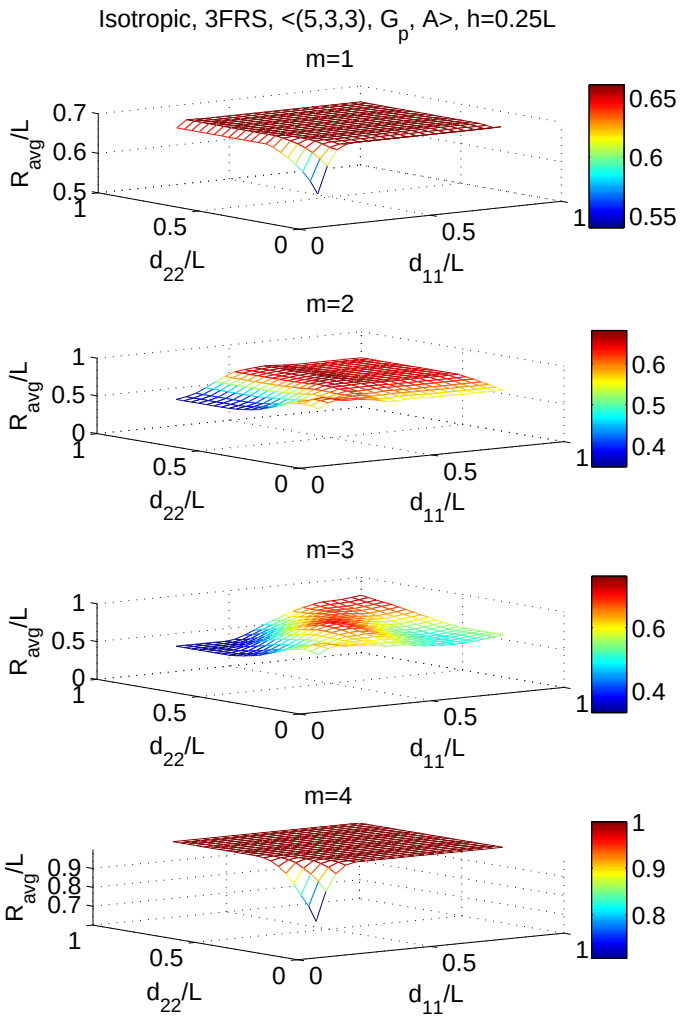

Figure A16: 3FRS Average availability with different decision depths with isotropic antennas, $\sigma = 0$, $h = 0.25L$

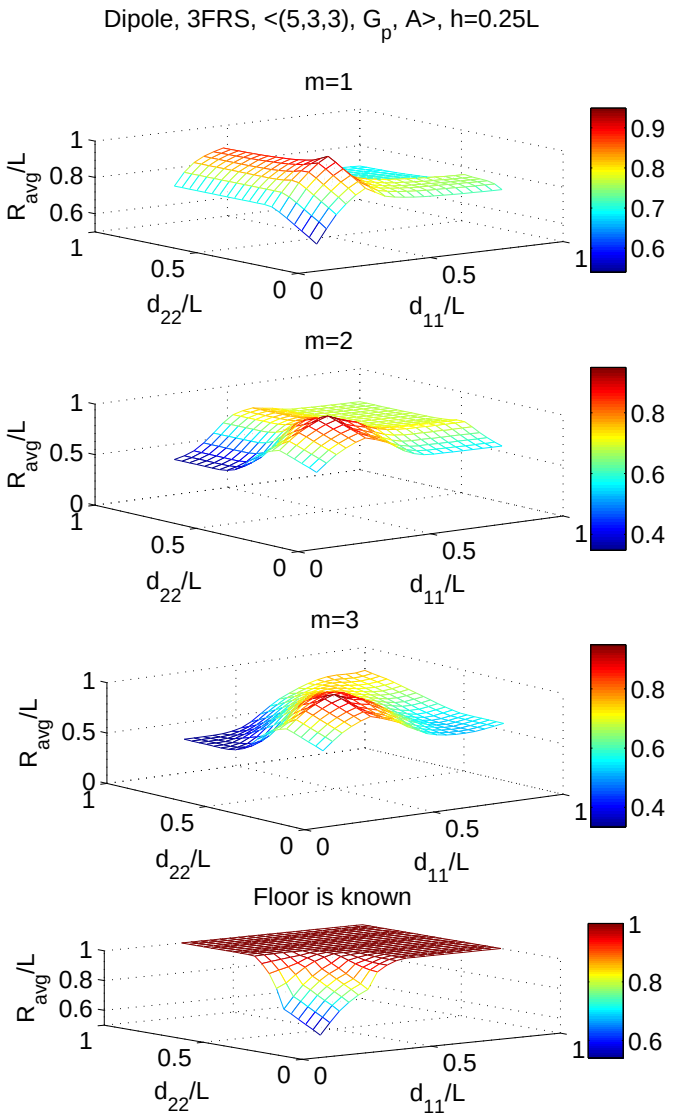

Figure A17: 3FRS Average availability with different decision depths with dipole antennas, $\sigma = 0$, $h = 0.25L$

A.3.1 Availability with MBFD

Average availability on three floors are shown in Figures A16 and A17 using isotropic and dipole antennas. We assume $h = 0.25L$. When $h < L/2$, floor determination algorithm causes the availability to decrease below 100% even under ideal channel conditions and monitoring range on current floor is large enough to cover the whole area on both of the floors. This was shown before in Figure A15. With dipole antennas in order for floor determination can be realized between multiple floors, $d > 2h$; therefore, we can neglect $d < 0.5L$ when $h = 0.25L$, and focus on where $d \geq 0.5L$. However, for the full coverage on all floors, $d \geq \sqrt{5}L/4$. Figures A18 and A19 show the availability that can be achieved with isotropic and dipole antennas on three floors.

Figures A20 and A21 show the probability of false negatives for three floors with isotropic and dipole antennas.

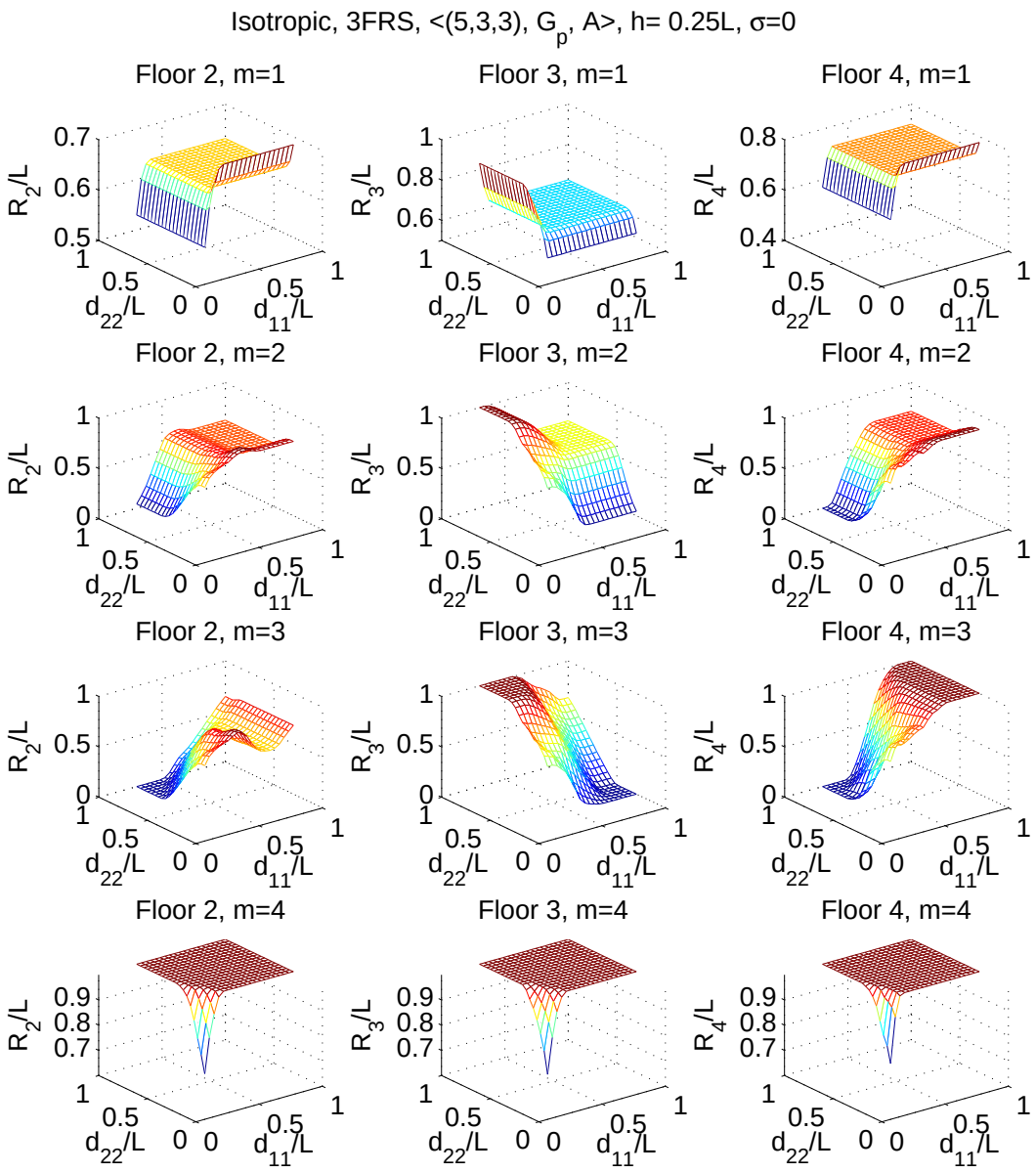

Figure A18: Availability on floors 2, 3 and 4 with MSAL-iso, $\sigma = 0$, $h = 0.25L$

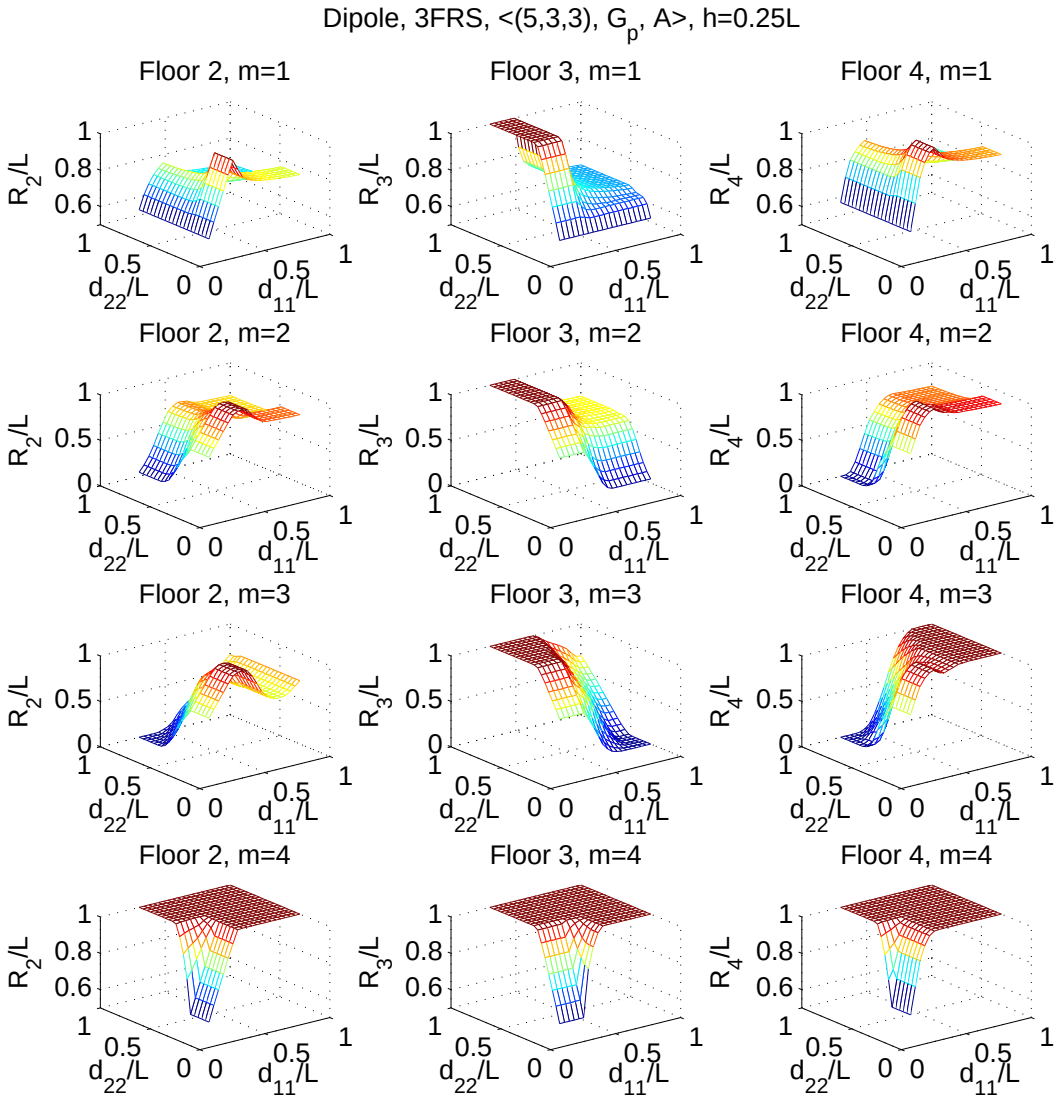

Figure A19: Availability on floors 2, 3 and 4 with MSAL-dpl, $\sigma = 0$, $h = 0.25L$

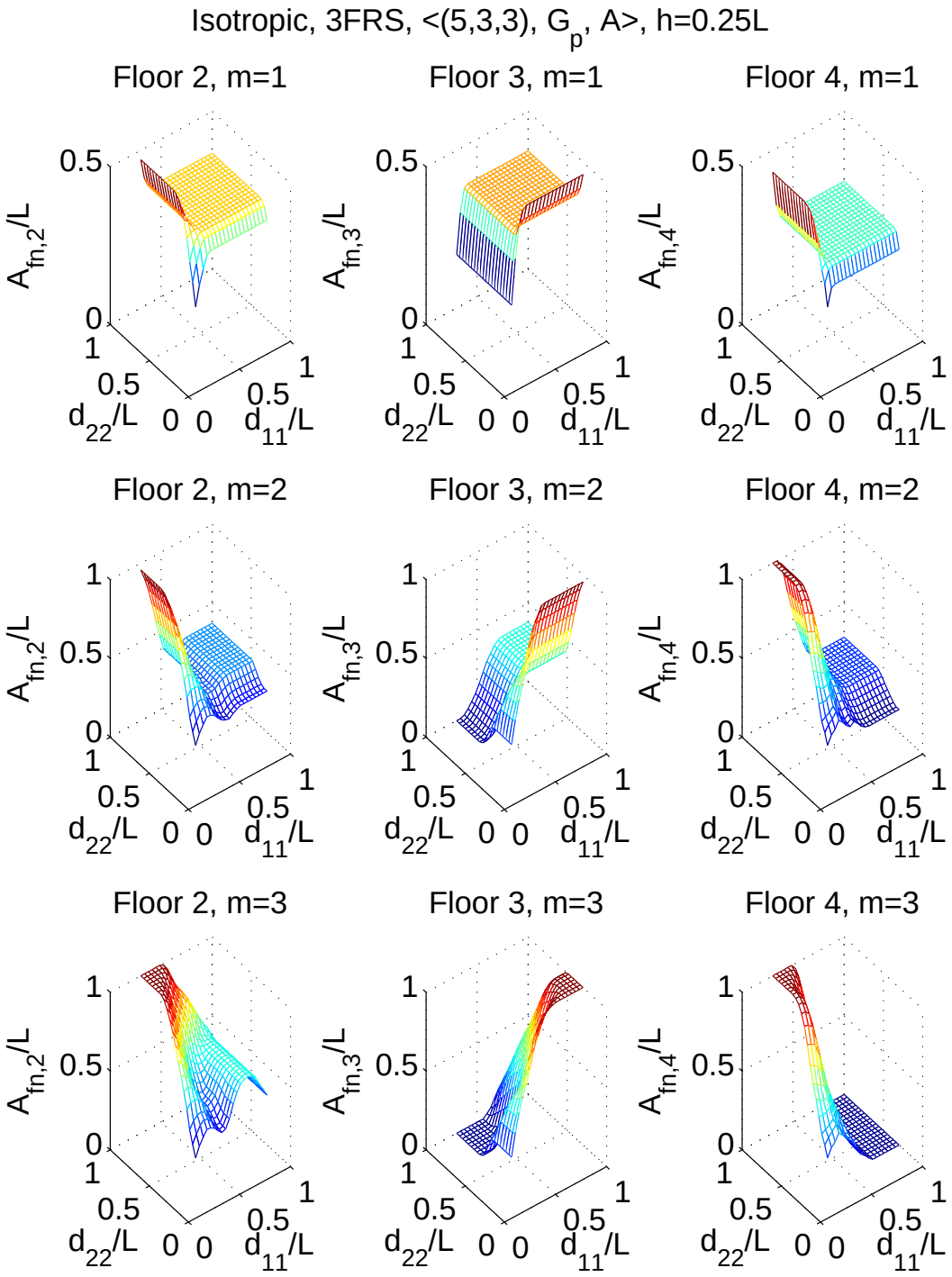

Figure A20: False negatives on floors 2, 3 and 4 with different decision depths with isotropic antennas, $\sigma = 0$, $h = 0.25L$

Figures A22 and A23 show the probability of false negatives for three floors on the average with isotropic and dipole antennas and with different decision depths used in floor determination.

A.3.2 Best Accuracy Conditions

Figures A24 and A25 show localization error with respect to two different monitoring ranges on floors 2,3 and 4.

Table A21 summarizes the minimum achievable errors and corresponding monitoring ranges with 3FRS scenario using MSAL-iso and MSAL-dpl under ideal conditions and $h = 0.25L$. Localization error when floor information is known is the upper bound for localization error that can be achieved with floor determination under ideal channel conditions. Because depending on floor determination algorithms performance, the availability of localization system increases or decreases. When availability decreases, the probability that a MN estimation increasing the average accuracy level gets smaller.

Figure A26 shows the accuracy with different floor heights under ideal channel conditions with 3FRS. With both MSAL and MCA, $m = 3$ is providing the best accuracy levels, while $m = 1$ has the worst accuracy levels. This is due to availability degradation with larger decision depth. Decision depth $m = 1$ provides the closest accuracy levels to the levels with known floor information.

With dipole antenna assumption, best achievable accuracy approaches to the same value for different decision depths at $h \geq 0.45L$ as shown in Figure A26. The reason can be explained as follows: When floor height is larger than a threshold, in order to use adjacent monitoring, monitoring ranges must be increased; however, increasing monitoring ranges may not be the best choice in accurate location estimation of MNs on the current floor. Therefore, when floor height is larger than this threshold, instead of using MSAL, SAL is used in location estimation. In this case, there is no need to floor determination algorithm and the percentage of false negatives is zero. Therefore, the number of MNs with estimated locations are the same.

Effect of log normal shadowing with $\sigma = 5$ on localization accuracy of 3FRS is given in

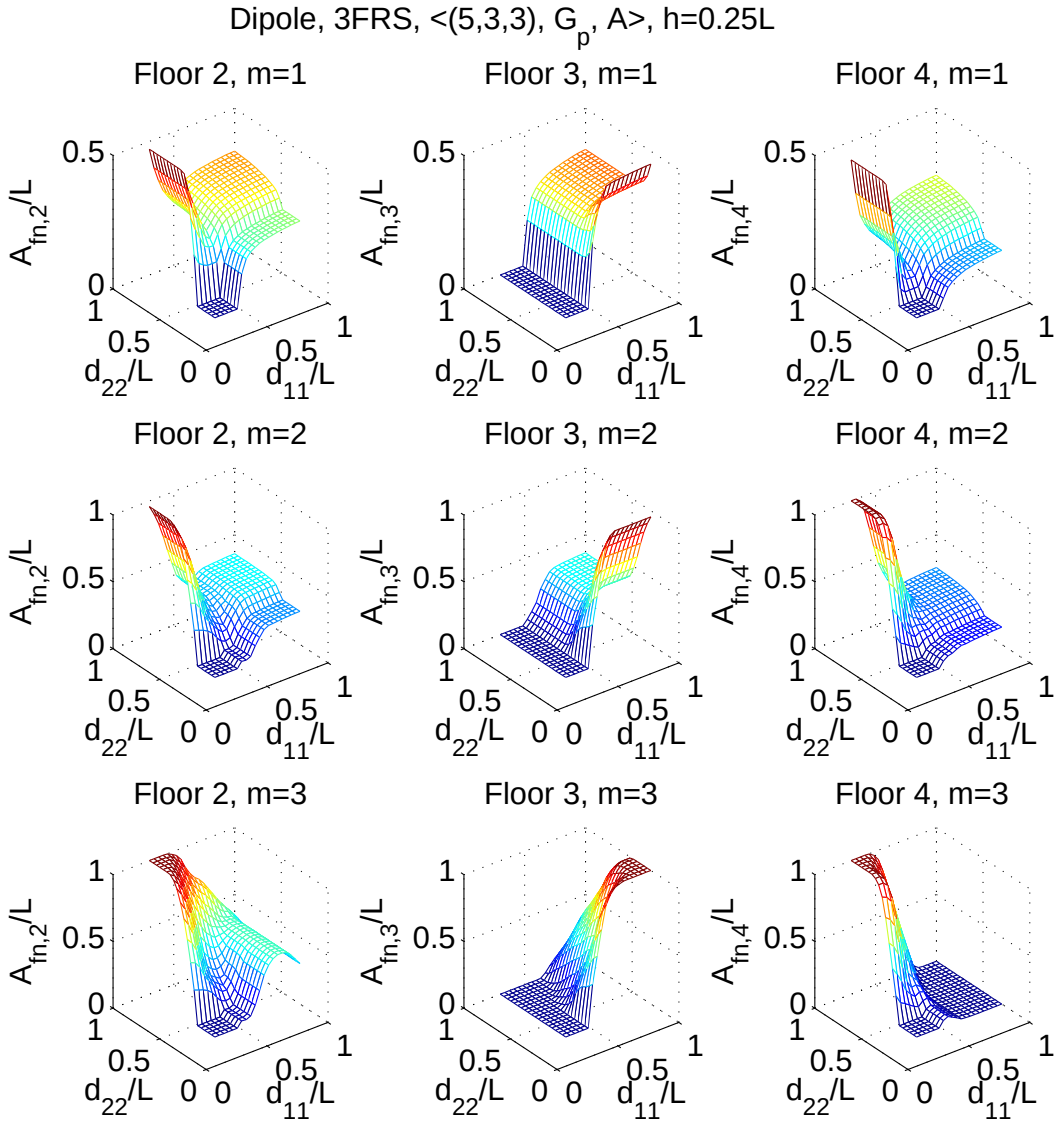

Figure A21: False negatives on floors 2, 3 and 4 with different decision depths with dipole antennas, $\sigma = 0$, $h = 0.25L$

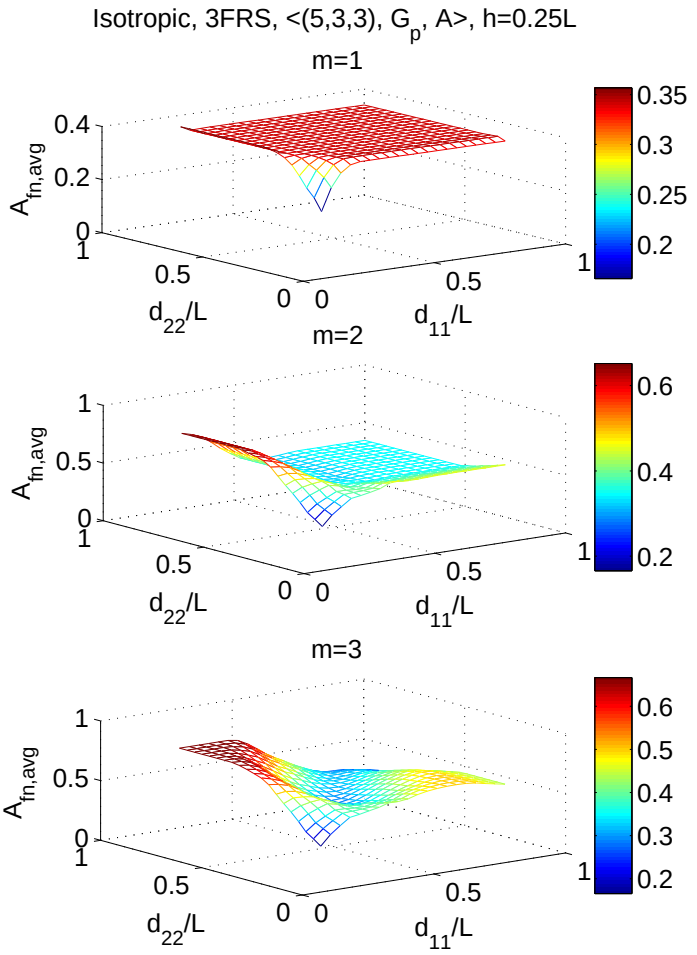

Figure A22: Average false negatives with different decision depths with isotropic antennas, $\sigma = 0$, $h = 0.25L$

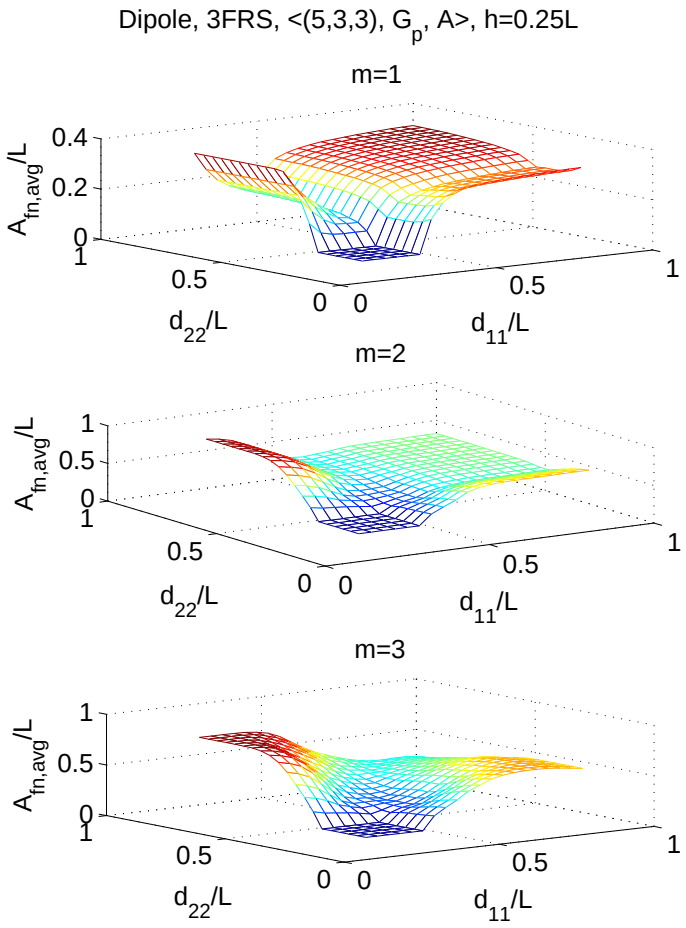

Figure A23: Average false negatives with different decision depths with dipole antennas, $\sigma = 0$, $h = 0.25L$

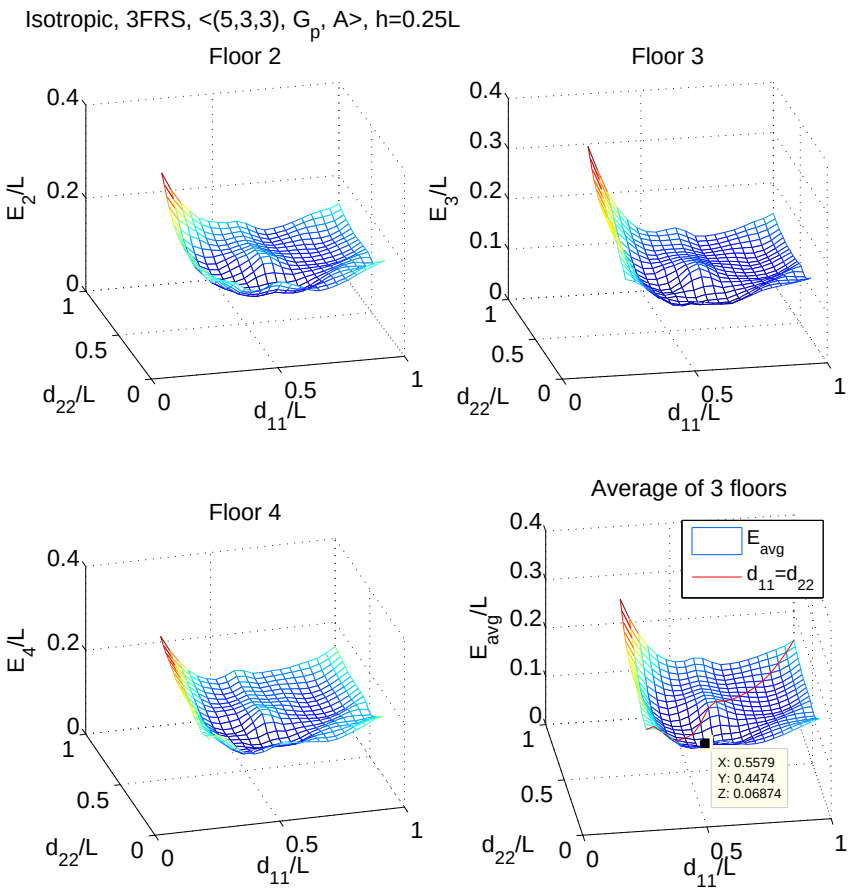

Figure A24: Localization errors on floors 2, 3 and 4 with MSAL-iso, $\sigma = 0$, $h = 0.25L$

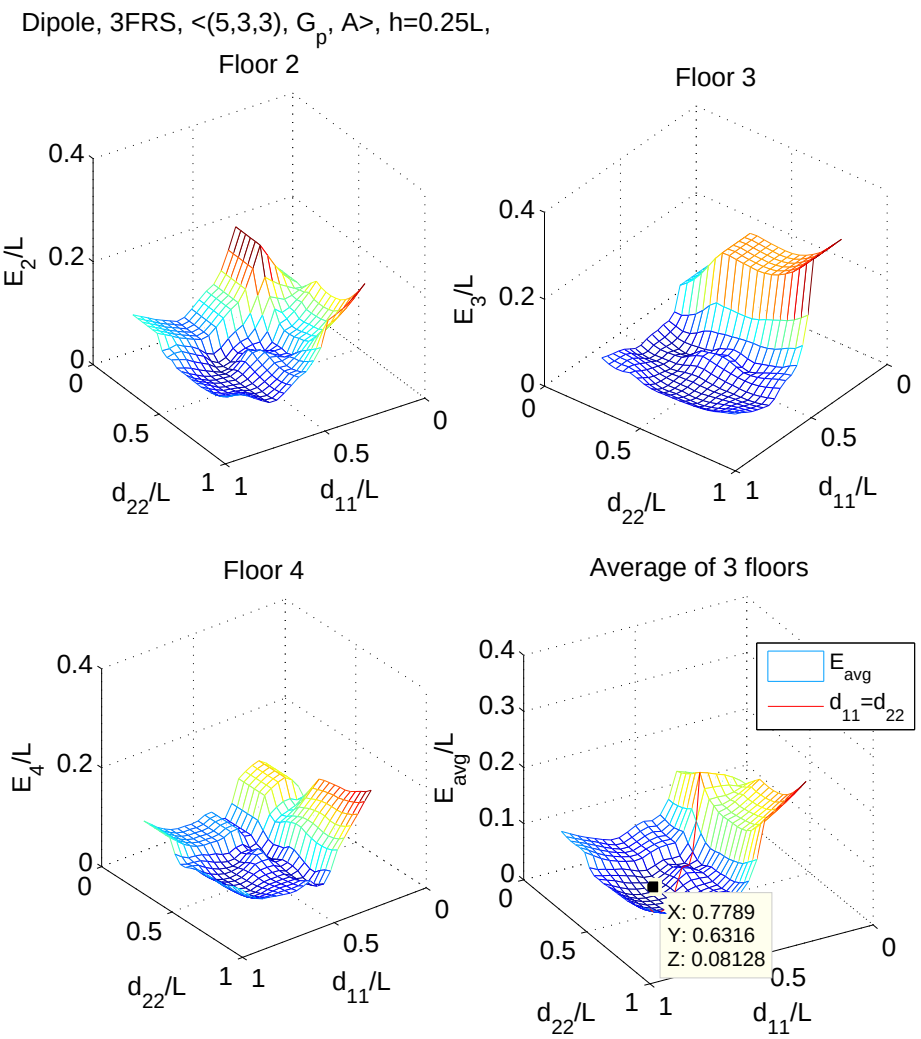

Figure A25: Localization errors on floors 2, 3 and 4 with MSAL-dpl, $\sigma = 0$, $h = 0.25L$

Table A21: Minimum localization error in 3FRS, $< (5,3,3), \mathbf{G}_p, A >$, $h = 0.25L$ $\sigma = 0$

	E_2/L	E_3/L	E_4/L	E^*_{avg}/L	$(d^*_{22}, d^*_{33}, d^*_{44})/L$
MSAL-iso	0.069	0.073	0.063	0.068	(0.55, 0.55, 0.44)
MSAL-dpl	0.078	0.09	0.075	0.0813	(0.78,0.78,0.63)

Figure A27 when different decision depths are used in FD. Our first observation is that the difference in accuracy caused by different decision depths is disappearing under shadowing with $\sigma = 5$ when MSAL+MCA or MCA is used. When MSAL is used in localization, $m = 1$ provides the closest accuracy to the accuracy with known floor information. Decision depths $m = 2$ and $m = 3$ causes worse availability therefore lower localization errors. Among the three localization algorithms MSAL performs best in terms of accuracy although availability with MSAL+MCA is higher. Error with $m = 3$ is the smallest due to higher percentage in false negatives at this decision depth.

A.3.3 Best Efficiency Conditions vs. Floor Height

As for best accuracy conditions, efficiency with known floor information constitutes the upper bound of efficiency that can be achieved with any floor determination algorithm under ideal conditions.

Figure A29 shows the maximum efficiency that can be achieved under ideal channel conditions with 3FRS. Maximum efficiency under ideal conditions is achieved by MSAL. As floor height increases, the maximum efficiency that can be achieved with FD approaches to maximum possible value that can be achieved assuming known floor information, therefore no false negatives. A_{fn} decreases as the floor height increases, this reflects as better availability and better efficiency. When $h > 0.5L$, availability stays at 100%; however accuracy gets affected by large floor height. Therefore, worse efficiency levels are seen. Although it is not possible to change the floor height, it looks like there is an optimum value for h that is

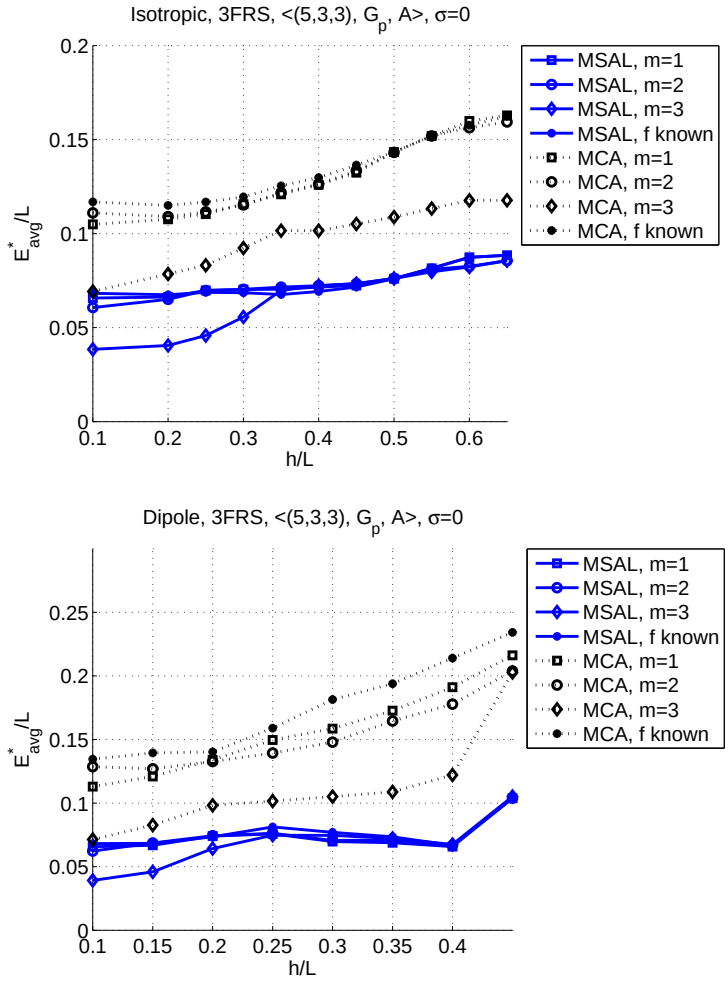

Figure A26: Minimum localization error with respect to h when $\sigma = 0$, 3FRS

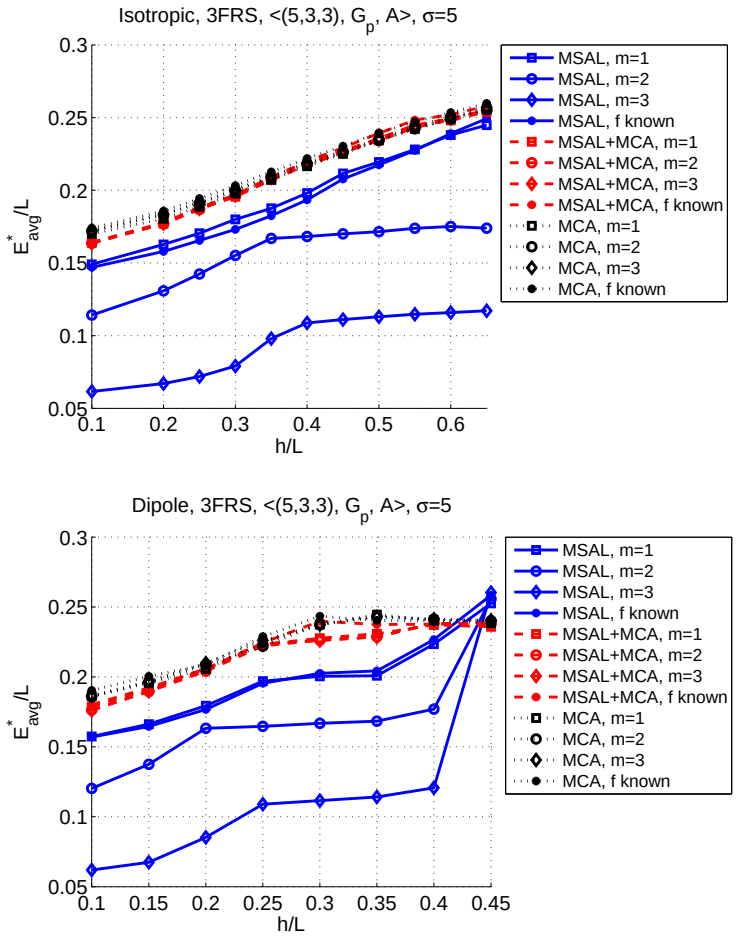

Figure A27: Minimum localization error with respect to h when $\sigma = 5$, 3FRS

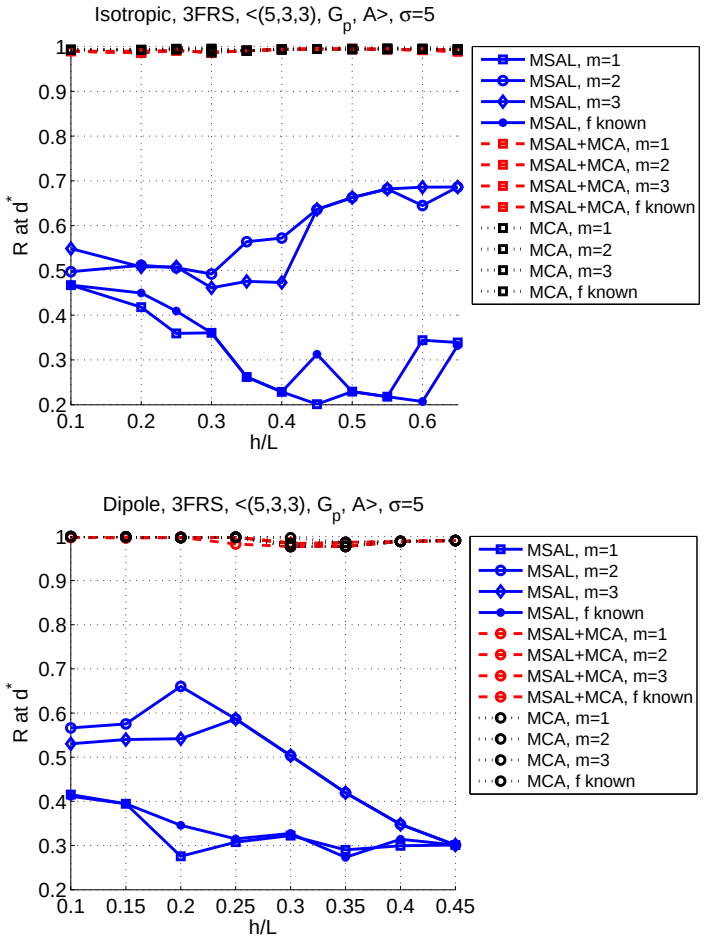

Figure A28: Availability at minimum localization error conditions with respect to h when $\sigma = 5$

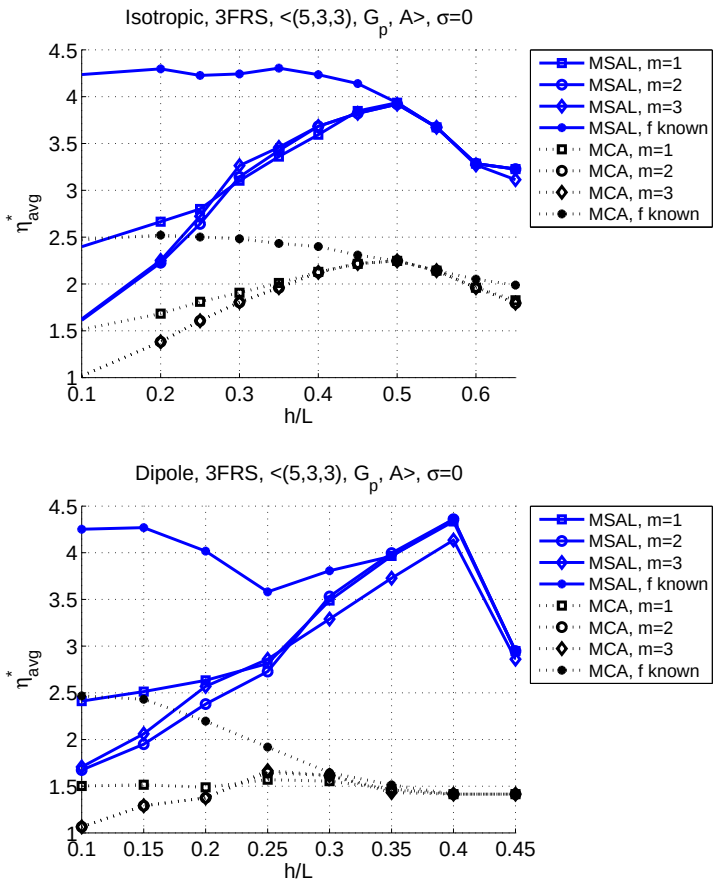

Figure A29: Maximum efficiency with respect to h when $\sigma = 0$, 3FRS

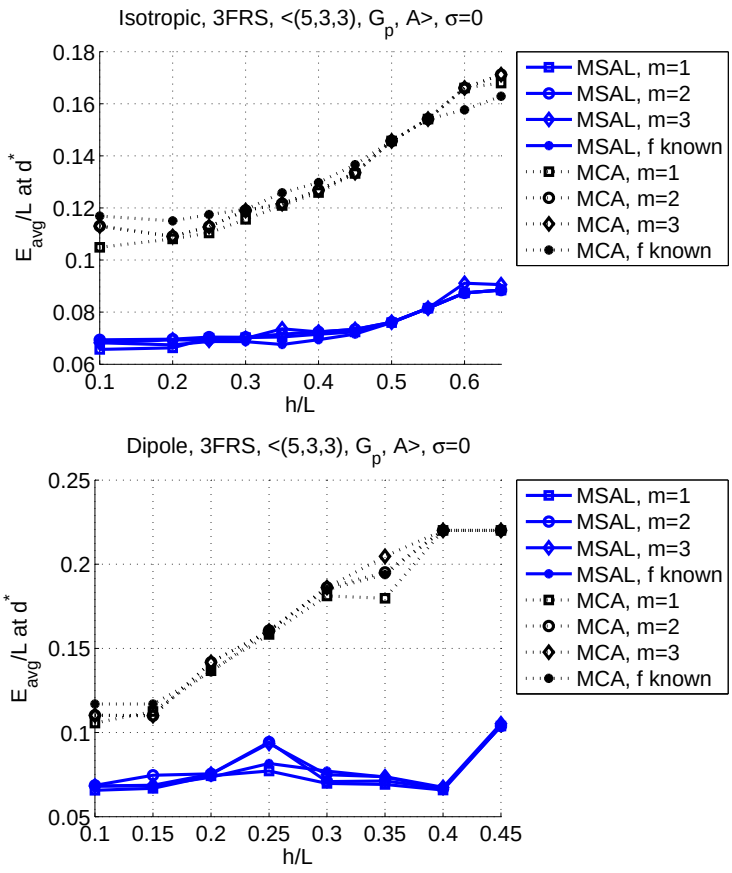

Figure A30: Average localization error at maximum efficiency conditions with respect to h when $\sigma = 0$

maximizing the efficiency. For isotropic coverage and dipole coverage, $h = 0.5L$ and $h = 0.4L$ are the best values. Accuracy at maximum efficiency levels with different decision depths are similar to each other.

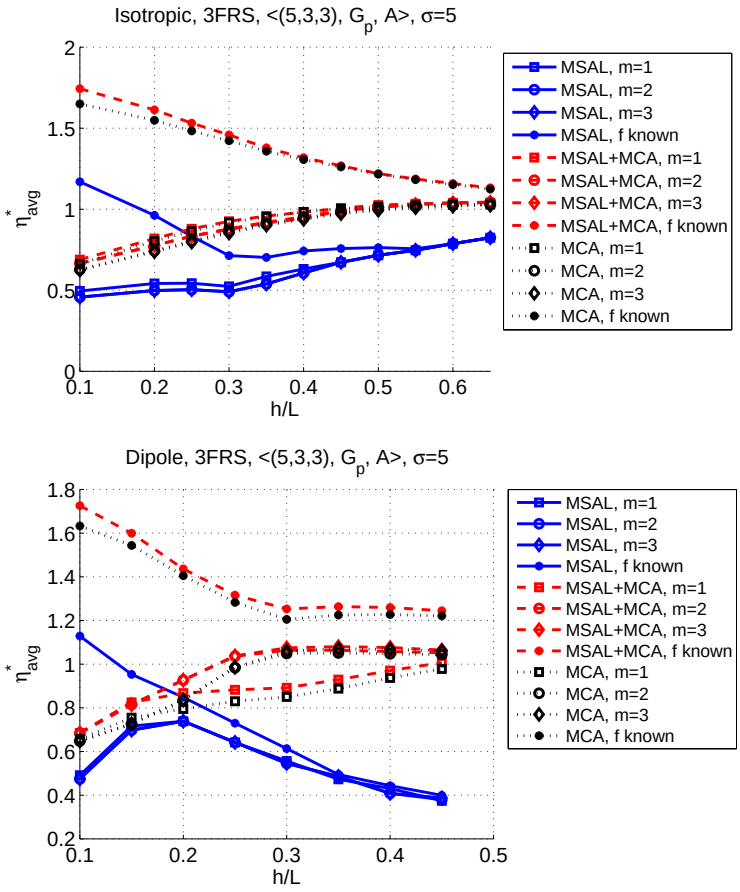

Figure A31: Maximum efficiency with respect to h when $\sigma = 5$, 3FRS

Figure A31 shows the maximum efficiency that can be achieved under shadowing with $\sigma = 5$ with 3FRS. Maximum efficiency is achieved by MSAL+MCA with known floor information. As floor height increases, maximum efficiency with floor determination approaches to maximum efficiency level as usage of adjacent coverages decreases gradually. Efficiency levels with different decision depths are similar to each other under this shadowing level. Assuming dipole antennas, decision depths $m = 2, 3$ provide the highest efficiency levels with MSAL+MCA.

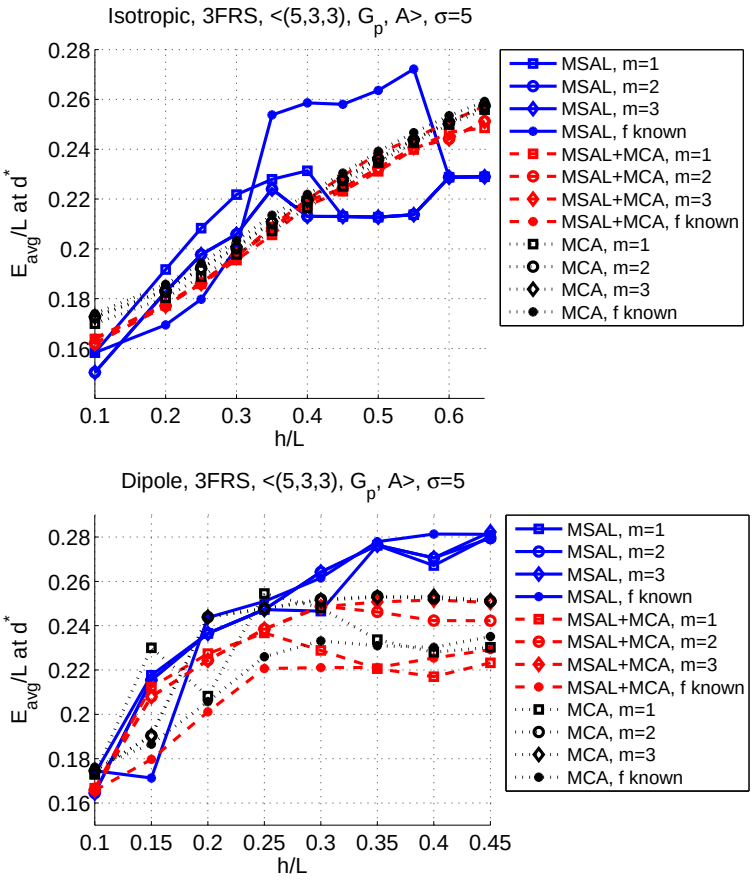

Figure A32: Average localization error at maximum efficiency with respect to h when $\sigma = 5$

A.3.4 Best Efficiency Conditions vs. Shadowing Effect

Finally, we show the effect of shadowing effect on maximum efficiency with 3FRS using floor determination algorithm with various decision depth values.

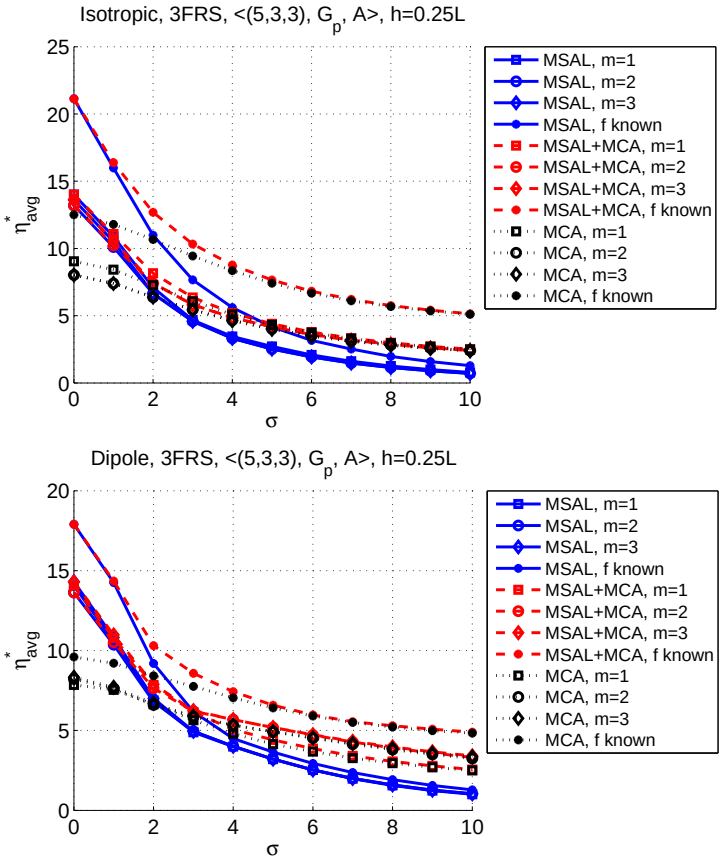

Figure A33: Maximum efficiency with respect to σ when $h = 0.25L$, 3FRS

Figures A33, A34 and A35 show the maximum efficiency, average localization error and availability possible at maximum efficiency conditions with respect to σ when $h = 0.25L$ with 3FRS. With isotropic antenna assumption, there is a significant degradation of 33% in maximum efficiency when floor determination is used. The decision depth effect on maximum efficiency is minimal. With dipole antenna assumption, floor determination causes less than 20% decrease in maximum achievable efficiency. Although there is not a significant difference in efficiency using different decision depths, $m = 2, 3$ provides highest efficiency

levels especially under high σ.

Figure A34 shows the average localization error at maximum efficiency conditions. At small σ values MSAL and MSAL+MCA perform similarly and outperform MCA, whereas at high σ values MSAL+MCA provides the smallest localization error. The observation is similar for both isotropic and dipole antenna assumptions.

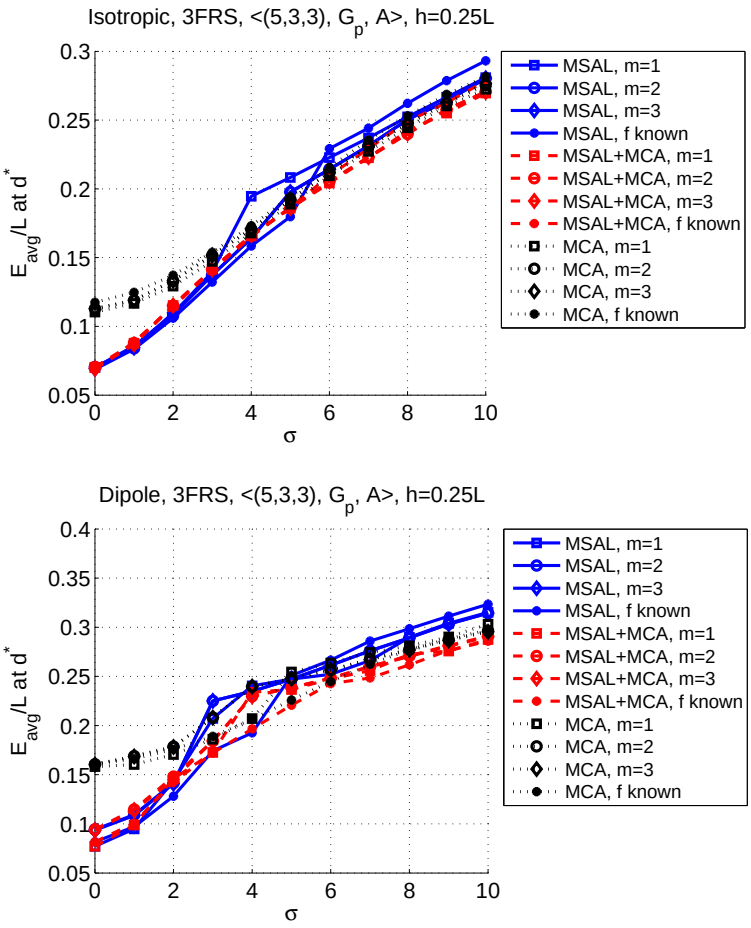

Figure A34: Average localization error at maximum efficiency conditions with respect to σ when $h = 0.25L$, 3FRS

Figure A35 shows the availability at maximum efficiency conditions. This figure also explains the degradation in maximum efficiency with isotropic antenna assumption. We observe a decrease in availability due to false negatives caused by floor determination algorithm.

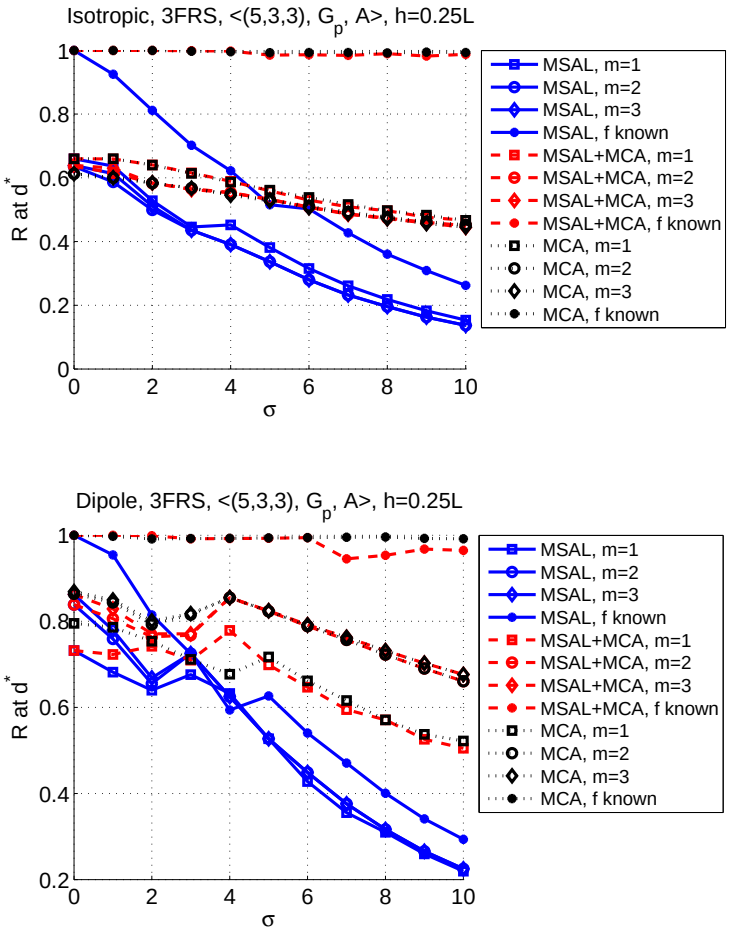

Figure A35: Availability at maximum efficiency conditions with respect to σ when $h = 0.25L$, 3FRS

A.3.5 Effect of Floor Attenuation Factor

A.3.5.1 Best Accuracy Conditions Figure A36 shows the minimum localization error with 3FRS under ideal channel conditions. When we compare this figure to Figure A26, we observe that the localization errors with different decision depths are approaching to the same value at a smaller floor height, i.e., $h \geq 0.35L$ and $h \geq 0.2L$ with isotropic and dipole antenna assumptions, respectively. Also, as expected, with floor attenuation factor localization error is higher than without floor attenuation. The reason is the narrower adjacent monitoring ranges because of larger path loss due to floor attenuation.

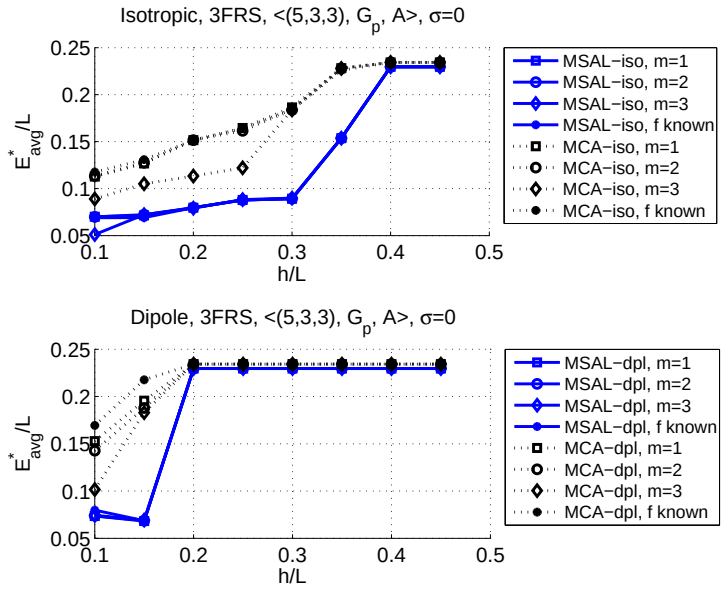

Figure A36: Minimum localization error with respect to h when $\sigma = 0$, 3FRS

Figure A37 demonstrates the same results from the false negatives point of view with 3FRS under best accuracy conditions. False negatives percentage approaches to zero as floor height increases. Effect of different decision depths used in floor determination algorithm vanishes when $h \geq 0.3L$ and $h \geq 0.15L$ as also observed in Figure A36. With floor height smaller than these values, effect of higher false negative percentage at $m = 2$ and $m = 3$ reflects as lower availability, and therefore lower localization error in Figure A36.

Similar conclusions can be drawn from Figure A39 showing the minimum localization

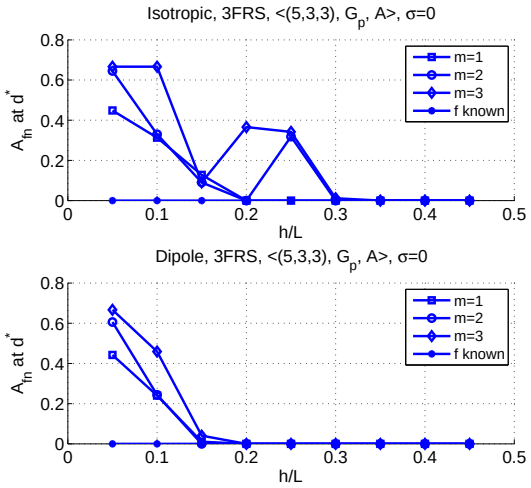

Figure A37: False negative percentage with respect to h when $\sigma = 0$, 3FRS

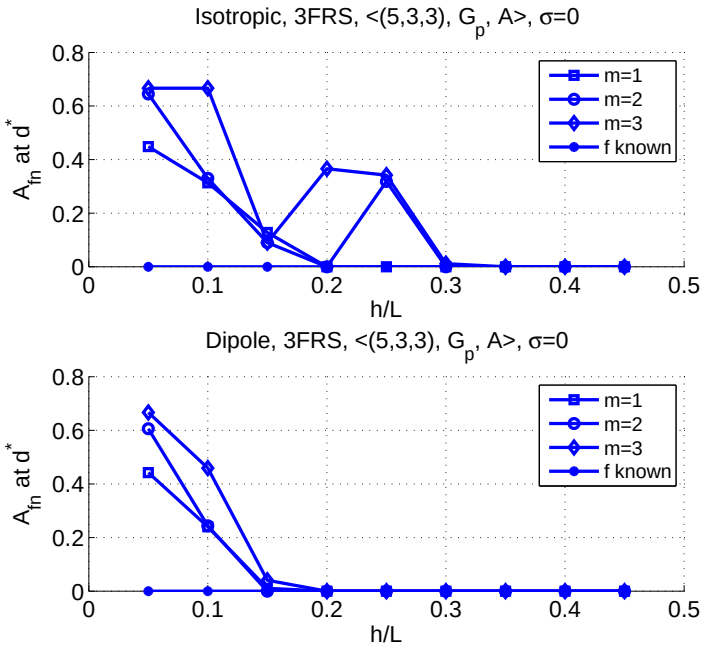

Figure A38: False negative percentage with respect to h when $\sigma = 0$, 3FRS

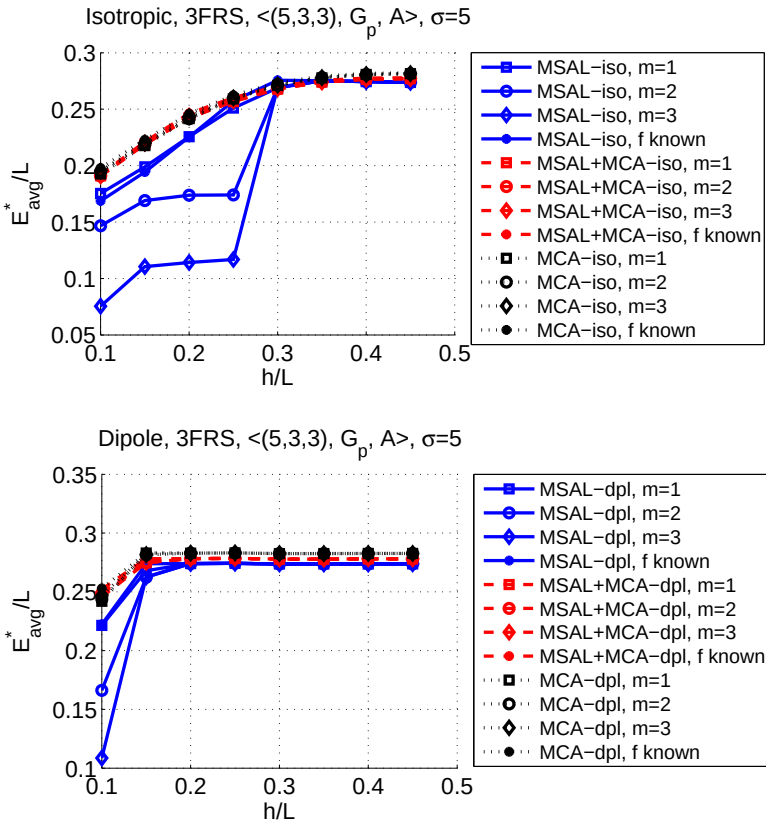

Figure A39: Minimum localization error with respect to h when $\sigma = 5$, 3FRS

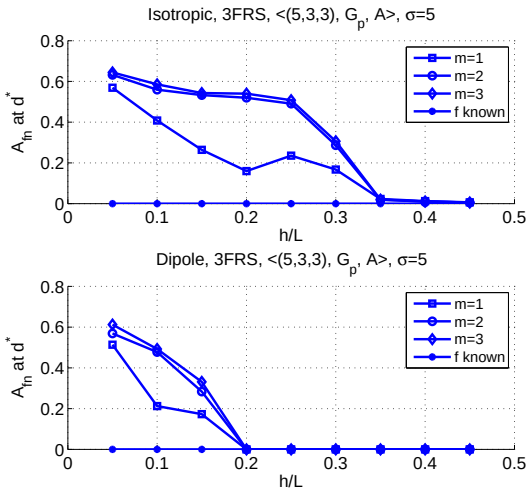

Figure A40: False negative percentage with respect to h when $\sigma = 5$, 3FRS

error under shadowing with $\sigma = 5$. As shown in Figure A40, false negatives percentage is higher under shadowing effects, and smallest with $m = 1$.

A.3.5.2 Best Efficiency Conditions vs. Floor Height The best efficiency with 3FRS that can be achieved under ideal channel conditions with floor attenuation factor are shown in Figure A41. When compared to Figure A29, we do not observe a significant decrease in efficiency values with isotropic antennas at small floor height. However, we do observe that at larger floor height, the efficiency with MSAL approaches to efficiency with MCA. This implies that at these floor height values, adjacent monitoring ranges approaches to zero, and SAL is used for localization. Since in SAL only current floor MoSs are used, the number of MoSs used in localization decreases, therefore accuracy gets worse as so efficiency. This is the reason why the efficiency approaches to the same values when $h \geq 0.4L$ and $h \geq 0.2L$ with isotropic and dipole antenna, respectively.

Best efficiency under shadowing with $\sigma = 5$ is shown in Figure A42. When compared to the case without FAF shown in Figure A31, we observe that efficiency with and without floor determination approaches to same values at a smaller floor height as expected. Also,

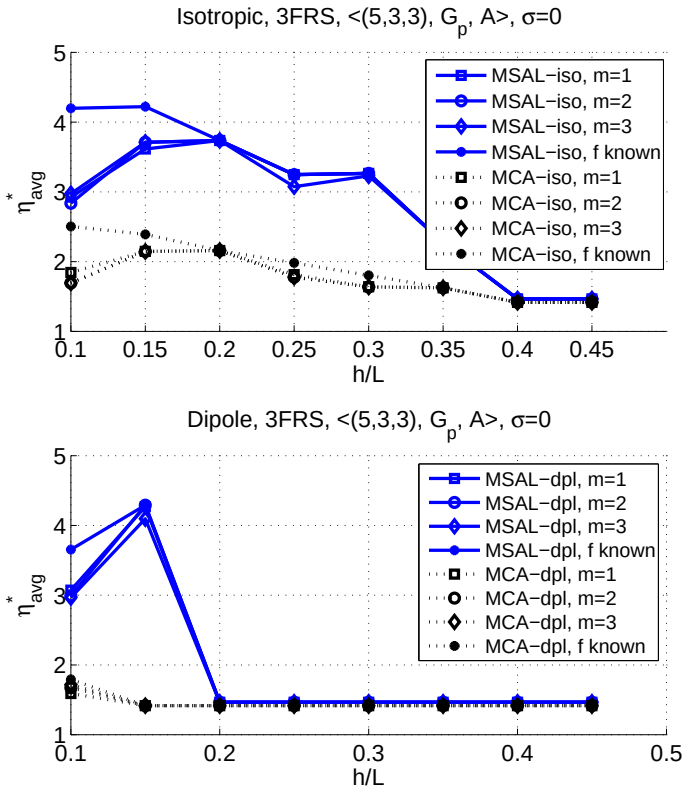

Figure A41: Maximum efficiency with respect to h when $\sigma = 0$, 3FRS

when floor heights is larger than these values, we observe similar results as SAL is used in localization instead of MSAL.

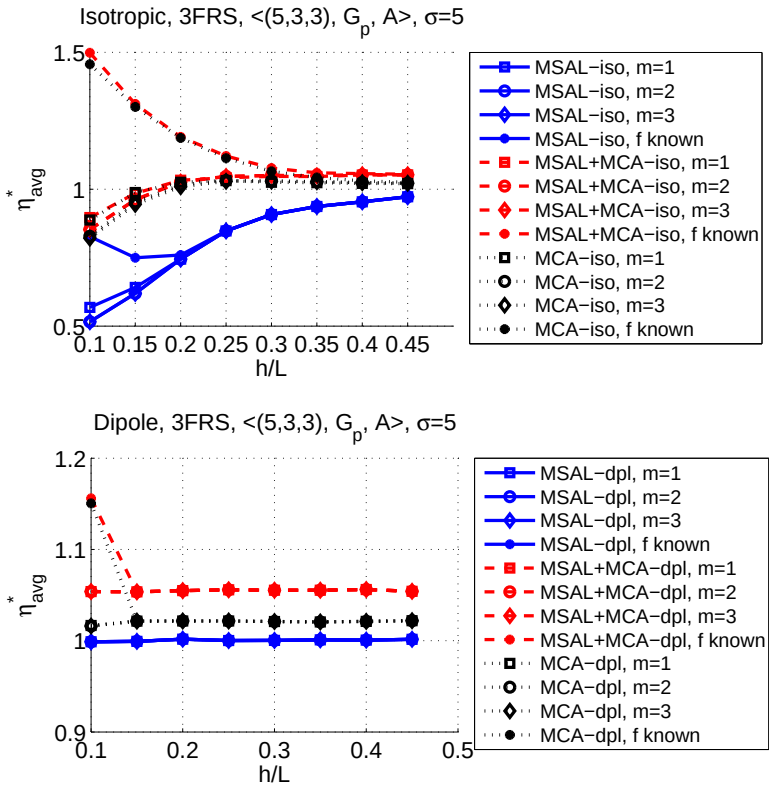

Figure A42: Maximum efficiency with respect to h when $\sigma = 5$, 3FRS

Figures A43 and A44 show the availability during best efficiency conditions under ideal conditions and shadowing with $\sigma = 5$. As also observed from the figures, we allow the coverage to be less than 100%. This is why at large floor height values, availability is not 100%. In addition, we observe higher availability at larger floor height, as the false negatives decrease at these values.

A.3.5.3 Best Efficiency Conditions vs. Shadowing Effect Lastly, we want to present the best efficiency conditions under shadowing with varying σ assuming $h = 0.25L$. Figures A45 and A46 show the best efficiency levels and localization accuracy available at these levels with and without floor determination. When compared to the case without FAF

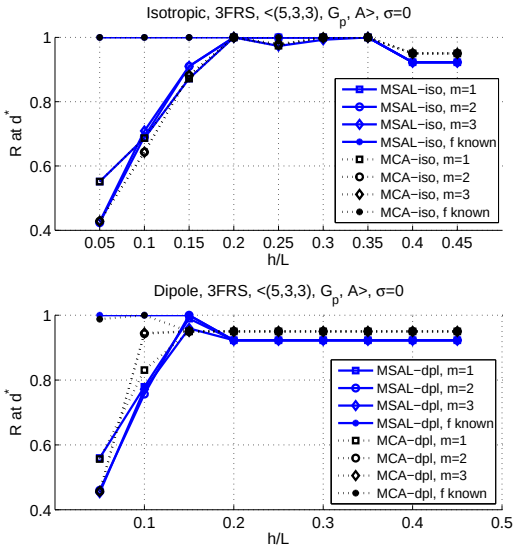

Figure A43: Availability with respect to h when $\sigma = 0$, 3FRS

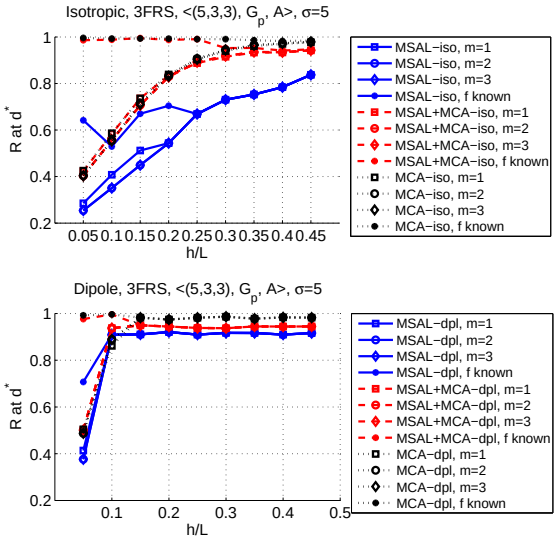

Figure A44: Availability with respect to h when $\sigma = 5$, 3FRS

shown in Figure A33, we observe a degradation in efficiency especially at lower σ values. At higher values, the difference vanishes. We observe with dipole antennas efficiency is the half of the efficiency achieved with isotropic antennas. In both cases, decision depth effect is insignificant, and MSAL+MCA outperforms other algorithms.

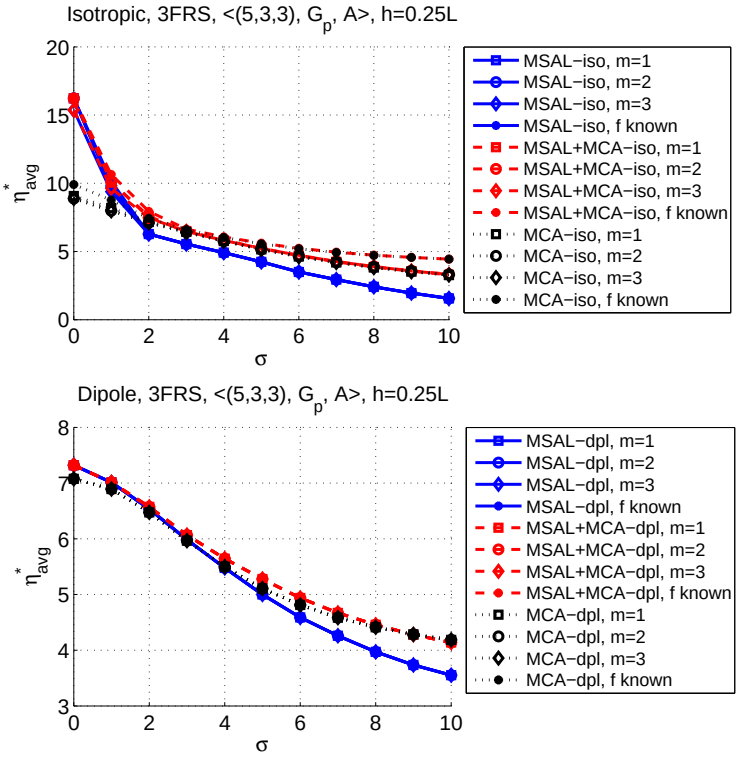

Figure A45: Maximum efficiency with respect to σ when $h = 0.25L$, 3FRS

From Figure A46, we observe MSAL+MCA and MSAL can provide best accuracy with dipole antennas for the whole range of σ values. And, with isotropic antennas, MSAL+MCA and MSAL perform similarly at low σ values, and as σ increases localization error of MSAL+MCA approaches to error with MCA. The reason is obviously the availability and false negatives percentage effects. False negative percentages with two types of antennas are shown in Figure A47. Availability is shown in Figure A48. Note that the sum of the availability with MSAL+MCA-iso and false negatives percentage with isotropic antennas makes 100%. However, we cannot draw the same conclusion for dipole antennas, as the percentage of false negatives is 0 at $h = 0.25L$. Availability is only affected by monitoring ranges and

shadowing.

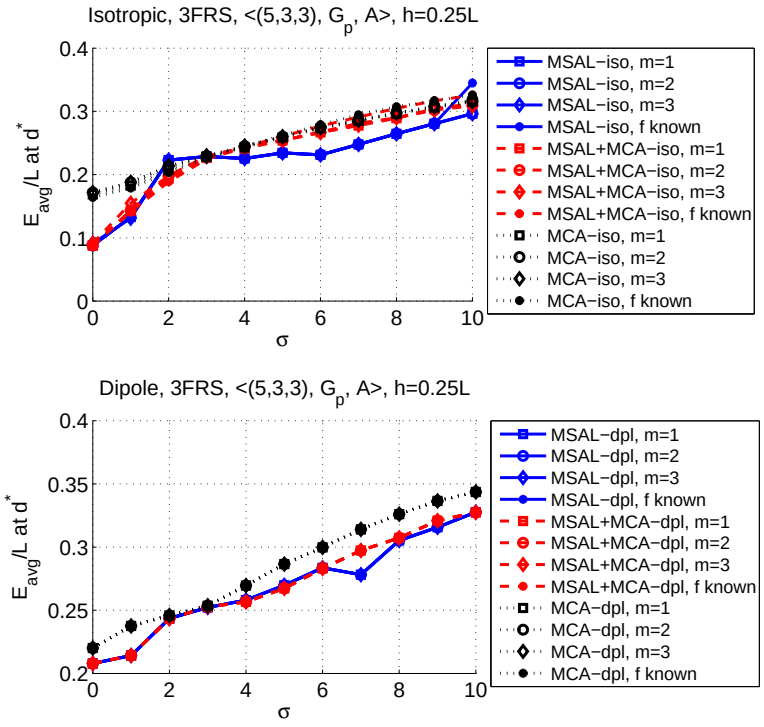

Figure A46: Localization error with respect to σ when $h = 0.25L$, 3FRS

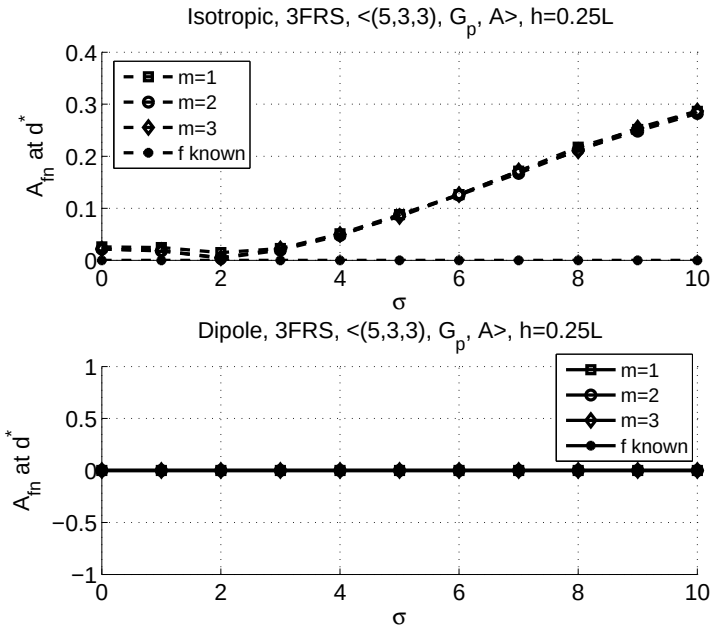

Figure A47: False negative percentage with respect to σ when $h = 0.25L$, 3FRS

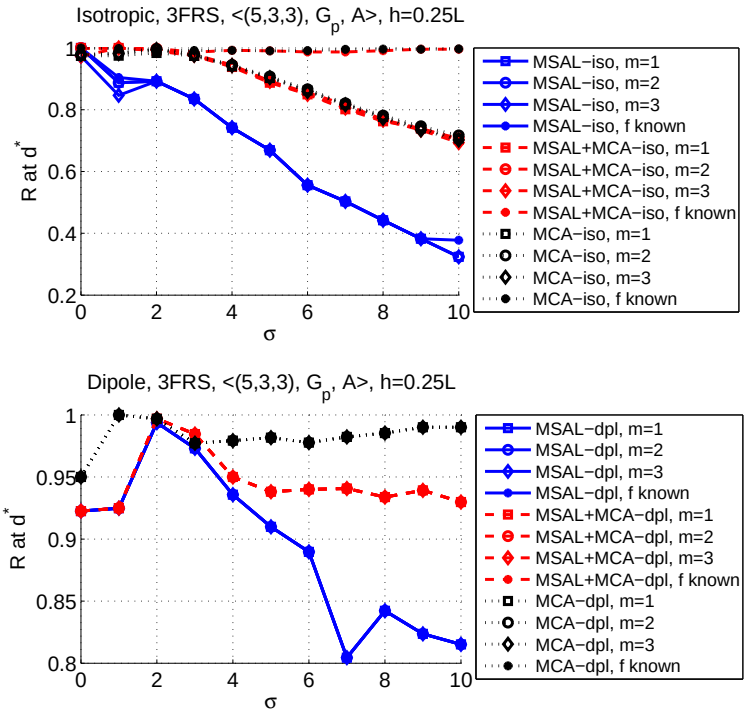

Figure A48: Availability with respect to σ when $h = 0.25L$, 3FRS

A.4 CASE STUDY

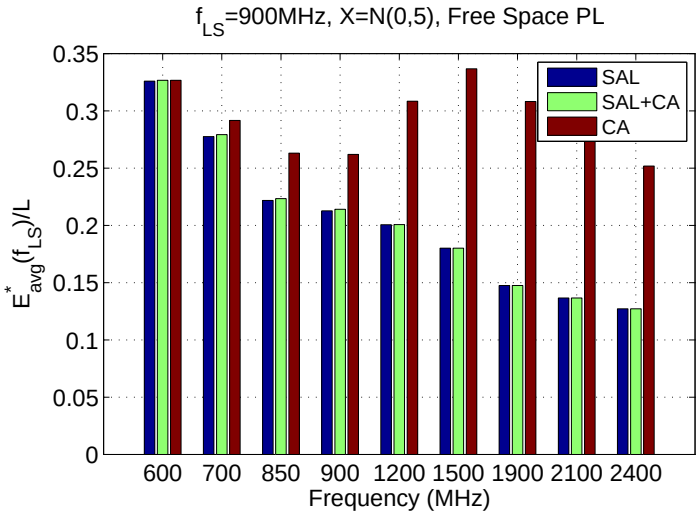

Figure A49: Performance when optimized for LS, $N = 4$ MoSs, free space, $f_{LS} = 900$MHz, $\sigma = 5$

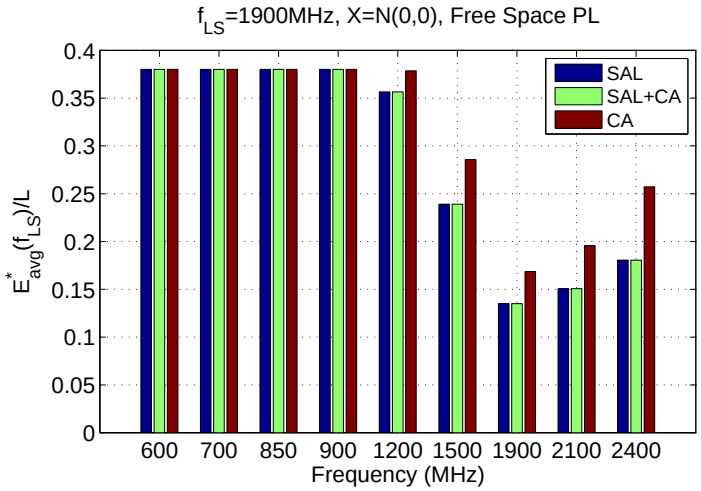

Figure A50: Performance when optimized for LS, $N = 4$ MoSs, free space, $f_{LS} = 1900$MHz, $\sigma = 5$

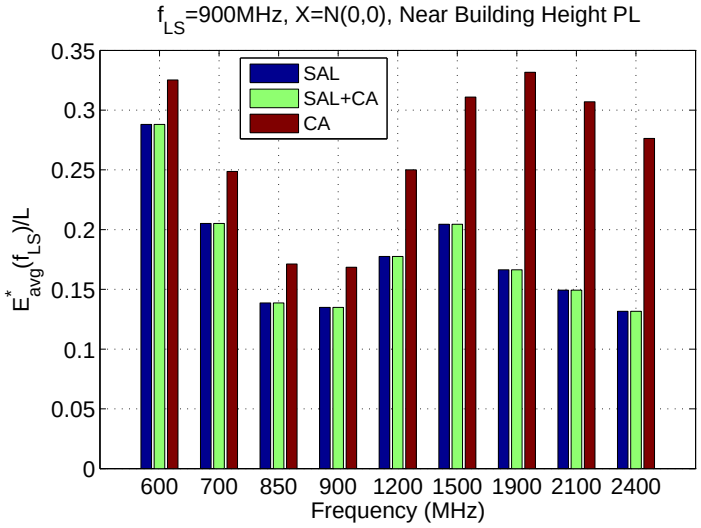

Figure A51: Performance when optimized for LS, building height, $f_{LS} = 900$MHz, $\sigma = 5$

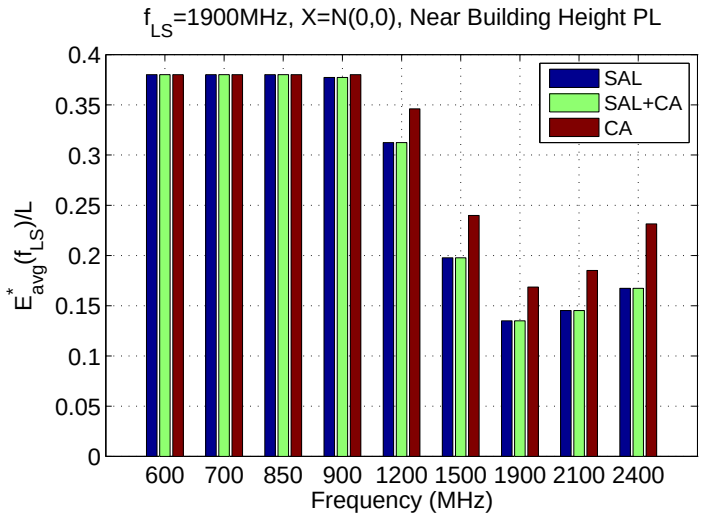

Figure A52: Performance when optimized for LS, building height, $f_{LS} = 1900$MHz, $\sigma = 5$

BIBLIOGRAPHY

[1] Y. Gwon and R. Jain, "Error characteristics and calibration-free techniques for wireless lan-based location estimation," in *MobiWac '04: Proceedings of the second international workshop on Mobility management & wireless access protocols*. New York, NY, USA: ACM, 2004, pp. 2–9.

[2] D. Madigan, E. Elnahrawy, R. P. Martin, J. Wen-Hua, P. Krishnan, and A. S. Krishnakumar, "Bayesian indoor positioning systems," *In Proc. of IEEE Infocom*, 2005.

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

[4] M. Hazas, J. Scott, and J. Krumm, "Location-aware computing comes of age," *Computer*, vol. 37, pp. 95–97, 2004.

[5] P. Mahonen, M. Petrova, and J. Riihijarvi, "Applications of topology information for cognitive radios and networks," *New Frontiers in Dynamic Spectrum Access Networks, 2007. DySPAN 2007. 2nd IEEE International Symposium on*, pp. 103 –114, apr. 2007.

[6] P. Enge, P.; Misra, "A survey on wireless ad hoc networks," *Proceedings of the IEEE*, vol. 87, pp. 3–15, Jan 1999.

[7] L. H. Hofmann-Wellenhof B. and C. J., *Global Positioning System: Theory and Practice*, 4th ed. Austria: Springer-Verlag, 1997.

[8] "http://www.gps.gov/."

[9] M. G. Rubinstein, I. M. Moraes, M. E. M. Campista, L. H. M. K. Costa, and O. C. M. B. Duarte, "A survey on wireless ad hoc networks," *Mobile and Wireless Communications Networks*, vol. 211, pp. 1–33, 2006.

[10] J. Hightower and G. Borriello, "Location systems for ubiquitous computing," *Computer*, vol. 34, no. 8, pp. 57–66, 2001.

[11] K. Muthukrishnan, M. Lijding, and P. Havinga, "P.: Towards smart surroundings: Enabling techniques and technologies for localization," in *In: Proceedings of the First*

International Workshop on Location and Context-Awareness (LoCA), SpringerVerlag, 2005.

[12] Y. Gu, A. Lo, and I. Niemegeers, "A survey of indoor positioning systems for wireless personal networks," *Communications Surveys Tutorials, IEEE*, vol. 11, no. 1, pp. 13 –32, 2009.

[13] A. Harter and A. Hopper, "A distributed location system for the active office," *IEEE Network*, vol. 8, no. 1, February 1994.

[14] D. Niculescu and B. Nath, "Vor base stations for indoor 802.11 positioning," in *MobiCom '04: Proceedings of the 10th annual international conference on Mobile computing and networking*. New York, NY, USA: ACM, 2004, pp. 58–69.

[15] D. Niculescu, "Positioning in ad hoc sensor networks," *Network, IEEE*, vol. 18, no. 4, pp. 24–29, 2004.

[16] N. Bulusu, J. Heidemann, and D. Estrin, "Gps-less low cost outdoor localization for very small devices," *IEEE Personal Communications Magazine*, vol. 7, no. 5, pp. 28–34, October 2000.

[17] J. J. Astrain, J. Villadangos, J. R. Garitagoitia, J. R. González de Mendívil, and V. Cholvi, "Fuzzy location and tracking on wireless networks," in *MobiWac '06: Proceedings of the 4th ACM international workshop on Mobility management and wireless access*. New York, NY, USA: ACM, 2006, pp. 84–91.

[18] N. Sastry, U. Shankar, and D. Wagner, "Secure verification of location claims," in *WiSe '03: Proceedings of the 2nd ACM workshop on Wireless security*. New York, NY, USA: ACM, 2003, pp. 1–10.

[19] S. Brands and D. Chaum, "Distance-bounding protocols," in *EUROCRYPT '93: Workshop on the theory and application of cryptographic techniques on Advances in cryptology*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 1994, pp. 344–359.

[20] S. Capkun, K. Rasmussen, M. Cagalj, and M. Srivastava, "Secure location verification with hidden and mobile base stations," *IEEE Transactions on Mobile Computing*, vol. 7, pp. 470–483, 2008.

[21] J. Hightower and G. Borriello, "A survey and taxonomy of location sensing systems for ubiquitous computing," University of Washington, Department of Computer Science and Engineering, Seattle, WA, UW CSE 01-08-03, August 2001.

[22] J. Webr and C. Lanzl, "Designing a positioning systems for finding things and people indoors," *IEEE Spectr.*, vol. 35, no. 9, pp. 71–78, 1998.

[23] R. Want, A. Hopper, V. Falc ao, and J. Gibbons, "The active badge location system," *ACM Trans. Inf. Syst.*, vol. 10, no. 1, pp. 91–102, 1992.

[24] L. M. Ni, Y. Liu, Y. C. Lau, and A. P. Patil, "Landmarc: indoor location sensing using active rfid," *Wirel. Netw.*, vol. 10, no. 6, pp. 701–710, 2004.

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

[26] H. Lim, L. C. Kung, J. C. Hou, and H. Luo, "Zero-configuration, robust indoor localization: theory and experimentation," *25th IEEE International Conference on Computer Communications Proceedings (INFOCOM)*, pp. 1–12, April 2006.

[27] M. Youssef, A. Youssef, C. Rieger, U. Shankar, and A. Agrawala, "Pinpoint: An asynchronous time-based location determination system," in *MobiSys '06: Proceedings of the 4th international conference on Mobile systems, applications and services*. New York, NY, USA: ACM, 2006, pp. 165–176.

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

[29] S. P. and J. Cadman, "A comparison of rf tag location products for real world applications," Ubisense, http://www.ubisense.net/technology/files/, Tech. Rep., 2004.

[30] A. Sayed, A. Tarighat, and N. Khajehnouri, "Network-based wireless location: challenges faced in developing techniques for accurate wireless location information," *Signal Processing Magazine, IEEE,*, vol. 22, no. 4, pp. 24–40, 2005.

[31] A. R. Jimenez, F. Seco, R. Ceres, and L. Calderon, "Absolute localization using active beacons: A survey and iai-csic contributions."

[32] T. Rappaport, *Wireless Communications: Principles and Practice*. Upper Saddle River, NJ, USA: Prentice Hall PTR, 2001.

[33] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. New York, NY, USA: Cambridge University Press, 2005.

[34] K. Whitehouse, C. Karlof, and D. Culler, "A practical evaluation of radio signal strength for ranging-based localization," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 11, no. 1, pp. 41–52, 2007.

[35] S. Y. Seidel and T. S. Rappaport, "914 mhz path loss prediction models for indoor wireless communications in multifloored buildings," *IEEE Transactions on Antennas and Propagation*, vol. 40, pp. 207–217, 1992.

[36] h. Bluetooth Technology.

[37] A. Aksu, J. Kabara, and M. B. Spring, "Reduction of location estimation error using neural networks," in *MELT '08: Proceedings of the first ACM international workshop*

on Mobile entity localization and tracking in GPS-less environments. New York, NY, USA: ACM, 2008, pp. 103–108.

[38] M. Azizyan and R. R. Choudhury, "Surroundsense: mobile phone localization using ambient sound and light," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 13, no. 1, pp. 69–72, 2009.

[39] E. Elnahrawy, J. Austin-Francisco, and R. P. Martin, "Adding angle of arrival modality to basic rss location management techniques," *Proc. of IEEE Int. Symposium on Wireless Pervasive Computing (ISWPC'07)*, February 2007.

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

[41] S. MacLean and S. Datta, "A lower bound on range-free node localization algorithms," *Wireless Communication Systems. 2008. ISWCS '08. IEEE International Symposium on*, pp. 628 –632, oct. 2008.

[42] M. Youssef, "Handling samples correlation in the horus system," in *In IEEE Infocom*, 2004, pp. 1023–1031.

[43] K. Kaemerungsi and P. Krishnamurthy, "Properties of indoor received signal strength for wlan location fingerprinting," *IEEE/ACM Mobiquitous*, August 2004.

[44] W. H. Wong, J. K. Ng, and W. M. Yeung, "Wireless lan positioning with mobile devices in a library environment," in *ICDCSW '05: Proceedings of the Third International Workshop on Mobile Distributed Computing*. Washington, DC, USA: IEEE Computer Society, 2005, pp. 633–636.

[45] K. Kaemarungsi and P. Krishnamurthy, "Modeling of indoor positioning systems based on location fingerprinting," in *In IEEE Infocom*, 2004.

[46] P. Krishnan, A. S. Krishnakumar, J. Wen-Hua, C. Mallows, and S. N. Gamt, "A system for lease: location estimation assisted by stationary emitters for indoor rf wireless networks," *INFOCOM*, vol. 2, pp. 1001–1011, March 2004.

[47] J. P. J. Krumm, "Minimizing calibration effort for an indoor 802.11 device location measurement system," Microsoft Research, Tech. Rep., 2003.

[48] M. Hagan, M.T. Menhaj, "A practical evaluation of radio signal strength for ranging-based localization," *Neural Networks, IEEE Transactions on*, vol. 5, no. 6, pp. 989 – 993, Nov 1994.

[49] R. Battiti, A. Villani, R. Villani, and T. L. Nhat, "Neural network models for intelligent networks: Deriving the location from signal patterns," in *in Proceedings of AINS2002, (UCLA*, 2002.

[50] K. Derr and M. Manic, "Wireless based object tracking based on neural networks," in *Industrial Electronics and Applications, 3rd IEEE Conference on*, Singapore, 2008, pp. 308–313.

[51] X. Ding, H. Li, F. Li, and J. Wu, "A novel infrastructure wlan locating method based on neural network," in *AINTEC '08: Proceedings of the 4th Asian Conference on Internet Engineering*. New York, NY, USA: ACM, 2008, pp. 47–55.

[52] E. A. Martnez, R. Cruz, and J. Favela, "Estimating user location in a WLAN using backpropagation neural networks," in *Advances in Artificial Intelligence IBERAMIA 2004*, ser. Lecture Notes in Computer Science, vol. 3315. Springer Berlin / Heidelberg, 2004, pp. 737–746. [Online]. Available: http://www.springerlink.com/content/84n764ka2e073avh/

[53] M. Heidari, F. Akgul, N. Alsindi, and K. Pahlavan, "Neural network assisted identification of the absence of direct path in indoor localization," *Global Telecommunications Conference, 2007. GLOBECOM '07. IEEE*, pp. 387 –392, Nov. 2007.

[54] http://www.makhfi.com/tutorial/introduction.htm, "Introduction to neural networks and knowledge modeling."

[55] M. Youssef, A. Agrawala, A. U. Shankar, and S. H. Noh, "A probabilistic clustering-based indoor location determination system," University of Maryland, College Park, UMIACS-TR 2002-30, March 2002.

[56] A. M. Ladd, K. E. Berkis, A. Rudys, L. E. Kavraki, and D. S. Wallach, "Robotics-based location sensing using wireless ethernet," *Wirel. Netw.*, vol. 11, no. 1-2, pp. 189–204, 2005.

[57] P. Bolliger, "Redpin - adaptive, zero-configuration indoor localization through user collaboration," in *MELT '08: Proceedings of the first ACM international workshop on Mobile entity localization and tracking in GPS-less environments*. New York, NY, USA: ACM, 2008, pp. 55–60.

[58] L. F. M. de Moraes and B. A. A. Nunes, "Calibration-free wlan location system based on dynamic mapping of signal strength," in *MobiWac '06: Proceedings of the 4th ACM international workshop on Mobility management and wireless access*. New York, NY, USA: ACM, 2006, pp. 92–99.

[59] S.-Y. Lau, T.-H. Lin, T.-Y. Huang, I.-H. Ng, and P. Huang, "A measurement study of zigbee-based indoor localization systems under rf interference," in *WINTECH '09: Proceedings of the 4th ACM international workshop on Experimental evaluation and characterization*. New York, NY, USA: ACM, 2009, pp. 35–42.

[60] N. Bulusu, J. Heidemann, and D. Estrin, "Adaptive beacon placement," *In Proc. of the 21st Int. Conf. on Distributed Computing Systems (ICDCS-21)*, pp. 16–19, April 2001.

[61] F. Benbadis, K. Obraczka, J. CortÈs, , and A. Brandwajn, "Exploring landmark placement strategies for topology-based localization in wireless sensor networks," *EURASIP J. Adv. Signal Process*, pp. 1–12, 2008.

[62] J. N. Ash and R. L. Moses, "On optimal anchor node placement in sensor localization by optimization of subspace principal angles," *In Proc. of theIEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2289–2292, April 2008.

[63] G. Han, D. Choi, and W. Lim, "Reference node placement and selection algorithm based on trilateration for indoor sensor networks," *Wirel. Commun. Mob. Comput*, vol. 9, no. 8, pp. 1017–1027, August 2009.

[64] I. T. Haque, I. Nikolaidis, and P. Gburzynski, "On the impact of node placement and profile point selection to indoor localization," *WMNC*, 2009.

[65] Y. Chen, J. austen Francisco, and R. P. Martin, "A practical approach to landmark deployment for indoor localization," in *in Proceedings of the Third Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON*, 2006, pp. 365–373.

[66] Y. Zheng, H. Wang, L. Wan, and X. Zhong, "A placement strategy for accurate toa localization algorithm," *In Proc. of the7th Annual Communication Networks and Services Research Conference*, pp. 166–170, May 2009.

[67] S. Meguerdichian, F. Koushanfar, G. Qu, and M. Potkonjak, "Exposure in wireless ad-hoc sensor networks," in *Proceedings of the 7th annual international conference on Mobile computing and networking*, ser. MobiCom '01. New York, NY, USA: ACM, 2001, pp. 139–150. [Online]. Available: http://doi.acm.org/10.1145/381677.381691

[68] Y. Zou and K. Chakrabarty, "Uncertainty-aware and coverage-oriented deployment for sensor networks," *J. Parallel Distrib. Comput.*, vol. 64, pp. 788–798, July 2004. [Online]. Available: http://portal.acm.org/citation.cfm?id=1034812.1034814

[69] A. Howard, M. J. Matarić, and S. G. S., "An incremental self-deployment algorithm for mobile sensor networks," *Auton. Robots*, vol. 13, pp. 113–126, September 2002. [Online]. Available: http://portal.acm.org/citation.cfm?id=591557.591931

[70] V. Otsason, A. Varshavsky, A. LaMarca, and E. de Lara, "Accurate gsm indoor localization," *In Proc. of the 7th International Conference on Ubiquitous Computing*, pp. 141–158, 2005.

[71] A. Varshavsky, A. LaMarca, J. Hightower, and de Lara, "E, "the skyloc floor localization system,"," *In Proceedings of the Fifth IEEE international Conference on Pervasive Computing and Communications*, pp. 125–134, 2007.

[72] W. ur Rehman, E. de Lara, , and S. Saroiu, "Cilos: A CDMA indoor localization system," *Proceedings of the 10th International Conference on Ubiquitous Computing (Ubicomp)*, 2008.

[73] J. Letchner, D. Fox, and A. LaMarca, *Large-Scale Localization from Wireless Signal Strength*. In Proc. the 20th National Conference on Artificial Intelligence, 2005.

[74] H.-Y. Wang, V. W. Zheng, J. Zhao, and Q. Yang, "Indoor localization in multi-floor environments with reduced effort," *The 8th Annual IEEE International Conference on Pervasive Computing and Communications*, 2010.

[75] R. Kershner, "The number of circles covering a set," *American Journal of Mathematics*, vol. 61, no. 3, pp. 665–671, 1939. [Online]. Available: http://www.jstor.org/stable/2371320

[76] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[77] E. Damosso, "Selection procedures for the choice of radio transmission technologies of the umts (umts 30.03)," *ETSI TR 101 112 v3.2.0*, April 1998.

[78] Q. Zhao and B. Sadler, "A survey of dynamic spectrum access," *Signal Processing Magazine, IEEE*, vol. 24, no. 3, pp. 79 –89, may. 2007.

[79] R. Tandra, M. Mishra, A. Sahai, S. Kristen, A. Woyach, P. Grover, H. P. Are, R. Tandra, S. Mubaraq, and M. A. Sahai, "Extended edition: What is a spectrum hole and what does it take to recognize one?" 2010. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.152.5665

[80] J. Riihijarvi and P. Mahonen, "Exploiting spatial statistics of primary and secondary users towards improved cognitive radio networks," *Cognitive Radio Oriented Wireless Networks and Communications, 2008. CrownCom 2008. 3rd International Conference on*, pp. 1 –7, may. 2008.

[81] A. Ghasemi and E. S. Sousa, "Spectrum sensing in cognitive radio networks: the cooperation-processing tradeoff," *Wirel. Commun. Mob. Comput.*, vol. 7, no. 9, pp. 1049–1060, 2007.

[82] J. Nasreddine, J. Riihijarvi, and P. Mahonen, "Location-based adaptive detection threshold for dynamic spectrum access," *New Frontiers in Dynamic Spectrum, 2010 IEEE Symposium on*, pp. 1 –10, apr. 2010.

[83] Y. Zhao, L. Morales, J. Gaeddert, K. Bae, J.-S. Um, and J. Reed, "Applying radio environment maps to cognitive wireless regional area networks," *New Frontiers in Dynamic Spectrum Access Networks, 2007. DySPAN 2007. 2nd IEEE International Symposium on*, pp. 115 –118, apr. 2007.

[84] J. Riihijarvi, P. Mahonen, M. Petrova, and V. Kolar, "Enhancing cognitive radios with spatial statistics: From radio environment maps to topology engine," *Cognitive Radio Oriented Wireless Networks and Communications, 2009. CROWNCOM '09. 4th International Conference on*, pp. 1 –6, jun. 2009.

[85] H. H. Xia, "A simplified analytical model for predicting path loss in urban and suburban environments," *IEEE Transactions on Vehicular Technology*, vol. 46, pp. 1040–1046, 1997.

[86] Wikipedia, "Trump world tower." [Online]. Available: http://en.wikipedia.org/wiki/Trump_World_Tower

[87] "Council on tall buildings and urban habitat." [Online]. Available: http://www.ctbuh.org/TallBuildings/HeightStatistics/HeightCalculator/tabid/1007/language/en-US/Default.aspx