

Automated Tuning of Wireless Mesh Networks for Smart Meter Applications

by

Parham Zarei

BASc, Bu-Ali Sina University, 2016

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Electrical & Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

December 2019

© Parham Zarei, 2019

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, a thesis entitled:

Automated Tuning of Wireless Mesh Networks for Smart Meter Applications

submitted by Parham Zarei in partial fulfillment of the requirements for

the degree of Master of Applied Science

in Electrical & Computer Engineering

Examining Committee:

David G Michelson

Supervisor

Vijay K Bhargava

Supervisory Committee Member

Mieszko Lis

Additional Examiner

Abstract

Many power utilities use wireless mesh networks to interconnect the smart meters that support monitoring, protection and optimization of the distribution grid. Network performance is critically dependent on the distribution of smart meters/mesh nodes and the path-loss over the links between nodes and degrades as the number of nearest neighbours seen by each node: 1) increases in regions of high node density (leading to mutual interference) or 2) decreases in regions of low node density (leading to reduced reliability). Although this number can be reduced or increased by adjusting the transmit power of existing nodes and/or adding additional relay nodes as appropriate, manual tuning is extremely labour intensive and automated tuning algorithms that support both functions have not been previously reported. This work contributes to the development of practical automated tuning algorithms for smart meter networks in three ways. First, we show how the accuracy of simple path-loss models that characterize the relationship between mean path-loss and distance by a linear regression line degrades as path length decreases and the degree of shadow fading increases when using: 1) the ordinary least square (OLS) approach when distance measurements are error-free and 2) the errors-in-variables (EIV) approach when distance measurements are corrupted by errors. The results allow researchers to assess the reliability of short-range data sets and determine when EIV should be used in place of OLS. Second, we use these insights to develop a measurement-based short-range power-law path-loss model applicable to the smart meter environment using massive amounts of data obtained from BC Hydro's multi-service grid network (MSGN). The result is much more reliable than previous works based on more limited data and, for the first time, reveals that the long-term temporal variability of each link follows a lognormal distribution with the standard deviation across all links in a given region itself following a lognormal distribution. Finally, we propose and demonstrate the first distributed and combined relay node placement-transmit power adjustment (RNP-TPA) algorithm for wireless mesh networks that reduces mutual interference in high density parts and improves connectivity in low density parts of the network.

Lay Summary

Many power utilities use wireless mesh networks to interconnect the smart meters that support monitoring, protection and optimization of the distribution grid. Network performance degrades as the number of nearest neighbours seen by each node increases in regions of high node density (leading to mutual interference) or decreases in regions of low node density (leading to reduced reliability). Although this number can be reduced or increased by adjusting the transmit power of existing nodes and/or adding additional relay nodes as appropriate, manual tuning is extremely labour intensive and automated tuning algorithms that support both functions do not yet exist. Here we demonstrate an automated tuning algorithm that adds relay nodes and adjusts the transmission power of smart meters as required to achieve a more optimal nearest neighbour distribution. We also present a path-loss model applicable to smart meter environments that is more comprehensive and statistically reliable than previous efforts.

Preface

This thesis was prepared by Parham Zarei under the direct supervision of Prof. David G. Michelson in the Radio Science Lab (RSL) at University of British Columbia, Vancouver campus. Parham conducted the initial literature survey, worked with Prof. Michelson in developing an overall plan for the thesis, and analyzed data and conducted simulations to generate results, figures and tables. Dr. Zahra Vali, a postdoctoral fellow in RSL, contributed thoughtful discussion for developing Chapters 2, 3 and 4. Sol Lancashire, Manager of BC Hydro Telecommunications Services, provided data from BC Hydro's smart meter network management system for developing the path-loss model of Chapter 3 and provided useful insights for analyzing the data.

Table of Contents

Abstract.....	iii
Lay Summary	iv
Preface.....	v
Table of Contents	vi
List of Tables	ix
List of Figures.....	x
List of Abbreviations	xii
Acknowledgements	xiv
Dedication	xv
Chapter 1: Introduction	1
1.1 Significance.....	1
1.2 Previous Work and its Limitations	3
1.3 Objectives	5
1.4 Approach.....	6
1.5 Outline.....	7
Chapter 2: Statistical Assessment in Estimation of Slope in Short-Range Propagation Path- loss Model	8
2.1 Introduction.....	8
2.2 Concept	11
2.2.1 Regression Line	11
2.2.2 Ordinary Least Square Method.....	12

2.2.3	Cases of Error	12
2.3	Methodology	13
2.3.1	Admissibility Conditions	15
2.3.2	Synthesized Dataset	16
2.4	Results.....	17
2.4.1	Performance of MoM vs OLS.....	17
2.4.2	Applicability of MoM.....	24
2.4.3	Effect of Shadow Fading on MoM	27
2.5	Discussion.....	30

Chapter 3: Long-Term Temporal and Spatial Statistical Assessment of Short-Range

Power-Law Path-loss Model in Sub-Urban Environments with BC Hydro Utility Network

Dataset.....	32	
3.1	Introduction.....	32
3.2	Concept	34
3.3	Methodology	36
3.3.1	Data Filtering	37
3.3.2	Fading Characteristics.....	39
3.3.3	Path-loss Estimation Method	41
3.4	Results.....	41
3.4.1	Statistical Assessment on the BC Hydro MSGN Dataset.....	41
3.4.2	Long-Term Temporal and Spatial Variability in Path-loss Parameters.....	43
3.4.3	Estimate the Parameters of Path-Loss Model	44
3.5	Discussion.....	47

Chapter 4: Combined Relay Node Placement-Transmit Power Adjustment Algorithm in Non-Battery Based Wireless Sensor Network.....	49
4.1 Introduction.....	49
4.2 Concepts.....	53
4.2.1 Number of Nearest Neighbors	53
4.2.2 RNP-TPA Algorithm	54
4.2.3 RNP Algorithm	56
4.2.4 TPA Algorithm	58
4.3 Methodology.....	59
4.3.1 Topology Irregularity.....	59
4.4 Results.....	61
4.4.1 Domain of Applicability	62
4.4.2 Evaluate Algorithm Performance	63
4.4.3 Effectiveness of the Algorithm.....	68
4.5 Discussion.....	72
Chapter 5: Conclusion and Future Work.....	74
5.1 Conclusion	74
5.2 Limitation and Future Work	78
References.....	79

List of Tables

Table 2.1. Mean and standard deviation of MoM slope over $\sigma\delta$ for $\beta_{det}=-2$	22
Table 2.2. Mean and standard deviation of MoM slope over $\sigma\delta$ for $\beta_{det}=-3$	23
Table 3.1. No. of communication links and nodes before filtering	39
Table 3.2. Communication link information.....	40
Table 3.3. Number of communication links	41
Table 3.4. Communication links path length statistics for each region	42
Table 3.5. Communication links RSSI statistics for each region	42
Table 3.6. Path-loss parameters	47
Table 4.1. Simulation parameters	61
Table 4.2. Applicability of RNP-TPA, $P_0 = 25$ dBm.....	63
Table 4.3. Applicability of RNP-TPA, $P_0 = 30$ dBm.....	63

List of Figures

Figure 1.1. BC Hydro smart grid application framework.....	3
Figure 2.1. β versus. $\sigma\delta$ for OLS and MoM, (a) Distance span= [25 50] m, (b) Distance span= [25 75] m, (c) Distance span= [25 125] m and (d) Distance span= [25 225] m.....	19
Figure 2.2. Mean value of slope versus. $\sigma\delta$ for (a) $\sigma\varepsilon=1$ and $\beta_{det}=-2$ and (b) $\sigma\varepsilon=9$ and $\beta_{det}=-2$	20
Figure 2.3. Mean value of slope versus. $\sigma\delta$ for (a) $\sigma\varepsilon=1$ and $\beta_{det}=-3$, and (b) $\sigma\varepsilon=9$ and $\beta_{det}=-3$	21
Figure 2.4. (a) $\beta_{det}=-1$ (b) $\beta_{det}=-2$ (c) $\beta_{det}=-3$ (d) $\beta_{det}=-4$ (e) $\beta_{det}=-5$ (f) threshold of method selector index versus slope.....	27
Figure 2.5. Distance span = [25 50] m. (a) $\beta_{det}=-1$ (b) $\beta_{det}=-3$ (c) $\beta_{det}=-5$	29
Figure 2.6. Distance Span = [25 75] m. (a) $\beta_{det}=-1$ (b) $\beta_{det}=-3$ (c) $\beta_{det}=-5$	30
Figure 3.1. Typical Path-loss Model.....	35
Figure 3.2. Communication links under 1 km and smart meter nodes in the region of Terrace ..	39
Figure 3.3. Distribution of RSSI data for each city	43
Figure 3.4. CDF of path length.....	43
Figure 3.5. Daily value of slope.....	44
Figure 3.6. Daily value of shadow fading.....	44
Figure 3.7. Path-loss model and residual histogram of (a) Richmond, (c) Kamloops, (e)Terrace, Distribution of long-term temporal standard deviation of (b) Richmond, (d) Kamloops, (f) Terrace.	46
Figure 4.1. Proposed RNP-TPA algorithm.....	55

Figure 4.2. Average power of nodes vs. node density for various λ and P_0 65

Figure 4.3. Number of added relay nodes vs. node density for various λ and P_0 67

Figure 4.4. NNN vs. node density for various λ and P_0 69

Figure 4.5. σ NNN vs. node density for various λ and P_0 71

List of Abbreviations

A-D test	Anderson-Darling test
CDF	Cumulative Distribution Function
D2D	Device to Device
D2I	Device to Infrastructure
EIV	Error in Variable
ETX	Expected Transmission Count
GIS	Geographic Information System
GPS	Global Positioning System
ICT	Information Communication Technologies
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IoT	Internet of Things
LQI	Link Quality Indicator
LOS	Line of Sight
MAC	Medium Access Control
ML	Maximum Likelihood
MoM	Method of Moments
MSGN	Multi-service Grid Network
NIST	National Institute of Standards and Technology
NLOS	Non-Line of Sight
NMS	Network Management System
NNN	Number of Nearest Neighbor

OLS	Ordinary Least Square Method
OSI	Open System Interconnection
PLC	Power Line Communication
RNP	Relay Node Placement
ROLL	Routing over Low-Power, Lossy network
RSSI	Received Signal Strength Indication
SUN	Smart Utility Network
TPA	Transmit Power Adjustment
UHF	Ultra High Frequency
ULLN	Urban Low-power, Lossy Network
VHF	Very High Frequency
WMN	Wireless Mesh Network
WSN	Wireless Sensor Network
5G	Fifth Generation
6LoWPAN	IPv6 in Low power Wireless Personal Area Network

Acknowledgements

I would like to thank Dr. David G. Michelson for all his support, help, and for continuing to push me to meet higher standards of work. I am grateful to have you as my supervisor. I have learned and continue to learn a lot from you. Working with you has truly been an experience that has shaped my future and my values.

I would also like to thank my colleagues at the Radio Science Lab, Dr. Zahra Vali, Hamed Noori, Ibrahim Sanad, Mostafa Samadzadeh, Esther Xu, Aidan Hughes, Muneeba Raja, Anmol Bhardwaj and Amith Khandakar. I could have not wished for a better lab-mates. Thank you to Sol Lancashire for all his hard work for collecting BC Hydro data and providing valuable technical consultations.

And thank you to my wonderful family. To my mom, thank you for always being there for me. To my dad, thank you for always believing me.

Dedication

I would like to dedicate this thesis to Behrooz Zarei, Minoos Sharifi and Aleah Wielinga for their endless support.

Chapter 1: Introduction

1.1 Significance

Electricity delivery system in power grids starts after producing electricity in power plants which comprises of three parts: Transmission, Distribution, and Retailing. Demand response, automated fault location, isolation, and service restoration are among the applications of electricity delivery systems. The structure of delivery system has been redesigned from early 2000s by several organizations such as IEEE, USA-NIST smart grid road map and Canada-Ontario smart grid initiatives [1], [2]. Smart grid is the modern architecture of electricity delivery system which can monitor, protect, and optimize the network by two-way power and information flow in electrical and communication networks among interconnected parts of the system [3]. Each organization indicated the need to establish and implement information communication technologies (ICT) that maintain efficient, reliable, and secure connectivity between parts of the electrical delivery system to provide quick and effective response in parts of smart grids.

The majority of ICT infrastructure and standards are developed for electricity supply, transmission, and distribution substation of electricity delivery systems [4], [5]. Smart meter networks are an essential part of the smart grid as they can expand the ICT infrastructure to the entire distribution grid system and end users. The main benefit of smart meters for customers are as follows: 1) increase the efficiency of billing as meter reading happens daily, 2) opportunity to take better control over energy and consequently billing by providing the capability to shift the usage from on-peak to off-peak, 3) monitor the usage of the household or business online via internet and eventually, 4) real-time understanding of service interruptions for power utility companies [6], 5) provide near real-time electric system performance data required to push the system closer to operating limits and introduce large dynamic loads presented by widespread

adoption of electric vehicles [7], [8]. The advantages of smart meter networks for distribution grids are: 1) increase the reliability of power networks in outage management services as smart meters provide near real time reads of loads, voltage, and power factors, 2) distribution automation equipment such as voltage regulators and control switches can react in near real time for the purposes of demand response , 3) identification of fault location and network isolation will be enhanced in time [9], [10], and finally 4) safe and reliable integration of emerging and dynamic electric vehicle loads [8], [11]. Figure 1.1 shows BC Hydro's smart grid application framework for the distribution and customers in electricity delivery systems.

BC Hydro has one of the largest U-LLNs in the world with minimum 2 million smart meters in its sensing field which cover 95 % of BC population. In the coverage fields of BC Hydro, Terrain morphology, vegetation density, urban architecture, and smart meter and collector's density varies from one region to another. BC Hydro is seeking to leverage their U-LLN system to optimize their smart grid [7], [12], [13].

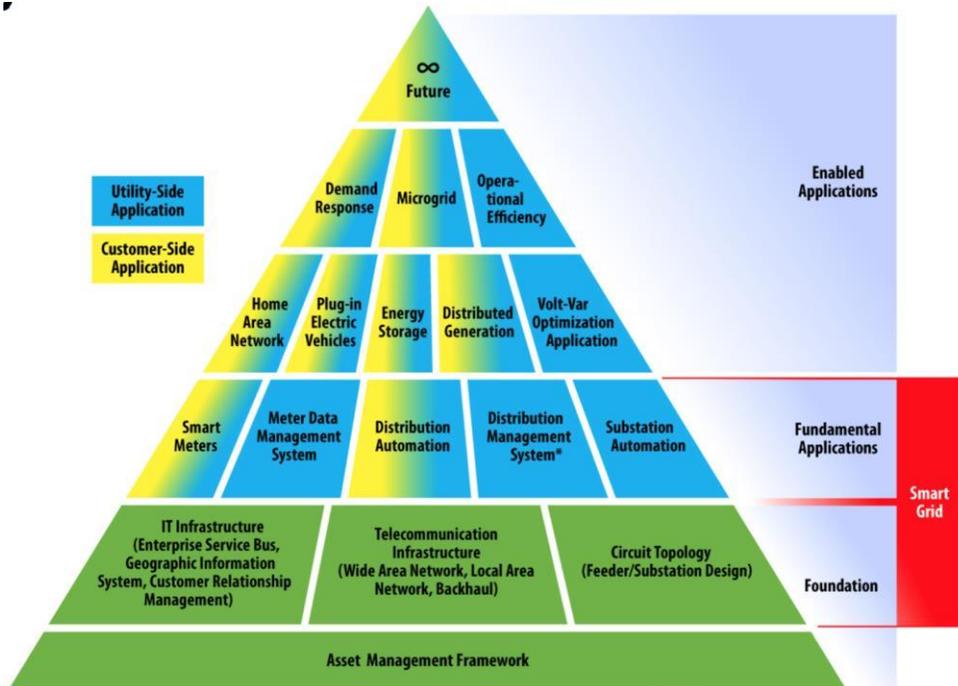


Figure 1.1. BC Hydro smart grid application framework

1.2 Previous Work and its Limitations

The number of end nodes in the distribution grids which needs to be connected with smart meter networks compared to the end nodes in transmission part of electrical delivery systems are several orders of magnitude higher. Therefore, the geographical regions that need to be covered for such networks are at the scale of a province and country. The implemented technologies for connectivity of smart meter networks can be classified based on the scalability, spectrum, data rate, coverage range, applications, and their limitations [14]. Wired power line communication (PLC), cellular network, and wireless mesh/ad hock networks are dominant technologies which are deployed to create smart meter networks [4].

PLC uses power lines to connect smart meters which makes it the most cost-effective technology. Data integrity and security are one of the disadvantages of the PLC as broadcasting

the messages in the power lines is in its nature. Additionally, the medium of PLC is harsh and noisy with a low bandwidth, therefore the quality of the signal is low which limits the number and type of devices that can be connected to the network. Cellular networks provide a secure, scalable, and high data rate connection between the smart meters. As the needs for smart meter communication are continuous, the capacity of cellular network needs to be shared with cellular network customers resulting in lower quality of service [15], [13].

Mesh network is a highly scalable, cost effective, and secure technology to connect a smart meter network. The ability to organize, configure, and heal itself while incorporating multi hop routing are among the its unique advantages which enables extending network coverage range and traffic balancing [16]. Physical and MAC layer of wireless mesh networks for utility networks are standardized with IEEE 802.15 smart utility network (SUN) task group 4g [17]. 802.15.4g low power radio supports: 1) operate from 700 MHz to 2.5 GHz in either licensed or license-exempted bands, 2) data rate between 40 to 1000 kbps 3) transmit small size IP packets without fragmentation and 4) coexistence with other systems in the same frequency bands. Interference between meters, fading of communication links which corresponds to connectivity malfunction between nodes, and network coverage due the incapability of meter density to cover the whole communication network are among the problems which exist in the physical layer of utility mesh networks [18].

Internet engineering task force (IETF) uses 6LowPAN (IPv6 in low power wireless personal area network) and ROLL (Routing over low power lossy links) task groups to standardize the other layer of OSI model [18]. Solutions have been proposed to address the problem of extending utility network coverage and reduce meter interference in the physical layer. key limitations of those researches are the idea that any issues in physical layer can be compensated in higher OSI layers, such as the network layer [19]. Some researches try to analyze the performance of urban low power

lossy network (U-LLN) metrics such as throughput and latency with different routing algorithms. Their key limitation was assuming the unrealistic model of wireless communication links [20], [21].

Performance of U-LLNs such as any other wireless mesh network is considerably affected by the wireless link environment. To fix the U-LLN's problems in the physical layer, a comprehensive path-loss model or any LQI metric such as expected transmission count (ETX) is needed to understand the performance of communication channels. As wireless links vary with time and environment, conducting measurements with lab grade instruments to understand the behavior of path-loss or LQI metric within the sensing field of a U-LLNs is neither cost effective nor time efficient. The only possible way to understand wireless links' performance in a network is to use the network performance data. Collecting network generated dataset is challenging due to all the constraints which other layers of OSI stack would impose to the physical layer and inaccuracy of the embedded measurement devices.

Once the performance of a communication channel is understood, interference between smart meters can be mitigated by adjusting the power of the nodes. Additionally, network coverage can be increased by finding the optimized location to add some nodes which can also help balancing its traffic. Given the scale of a U-LLN, manual tuning and replacement of smart meters require significant time and manpower. Therefore, a comprehensive distributed power adjustment and node placement algorithm is needed.

1.3 Objectives

In this work, our three main objectives are:

1. Propose and demonstrate combined and distributed relay node placement-transmit power adjustment (RNP-TPA) algorithm for wireless mesh networks that reduces mutual

interference in high density parts and improves connectivity in low density parts of the network.

2. Develop a measurement-based short-range power-law path-loss model which incorporates the spatial and long-term temporal variation of communication channels in various regions based on BC Hydro's MSGN dataset.
3. Demonstrate the slope estimation accuracy of simple path-loss model by linear regression line in short range links with shadow fading and distance error when using: 1) OLS and 2) EIV approaches.

1.4 Approach

Based on the mentioned objectives, the approach to tune BC Hydro's U-LLN is as follows:

The underlying statistical performance of simple path-loss model with regression line especially for short range links is investigated. The accuracy of path-loss estimation with OLS and EIV methods is investigated in the presence of distance error. Given the scale of distance error and shadow fading, a decision-making metric is proposed to determine when EIV should be used in place of OLS .

A statistical assessment is performed over BC Hydro MSGN dataset captured over three months in three different regions. The scale of distance error is measured, and the decision-making algorithm is applied to find the most accurate way to find path-loss parameters in the dataset. Finally, a comprehensive path-loss model which incorporates the long-term temporal and spatial variation of wireless links is proposed.

A combined, distributed RNP-TPA algorithm is proposed for wireless sensor networks with any random topology and various node densities. The algorithm performance is analyzed based on the reduction in smart meters' power and number of added nodes in the network.

1.5 Outline

The remainder of this thesis developed as follows: The underlying statistics in path-loss modeling is investigated in Chapter 2. A comprehensive path-loss model which incorporates long-term temporal and spatial variation of wireless links is proposed in Chapter 3. Combined relay node placement-transmit power adjustment algorithm is proposed in Chapter 4. Finally, conclusion and future works is presented in Chapter 5.

Chapter 2: Statistical Assessment in Estimation of Slope in Short-Range Propagation Path-loss Model

2.1 Introduction

Short-range wireless communications have drawn great attentions in recent years. With the advent of IoT, low power short range wireless technologies such as Bluetooth, ZigBee and 6LoWPAN have become more widespread [22], [23]. Accordingly, estimating the path-loss model of any wireless propagation channel model is of paramount importance. The importance of path-loss can be summarized in predicting link reliability and mutual interference. Log-distance path-loss model pursues a power-law relationship as follows:

$$PL(d)=PL(d_0)+10\beta \log_{10} \frac{d}{d_0} +X_{\sigma} \quad (1)$$

where $PL(d)$ is the path-loss at a distance d expressed in dB, $PL(d_0)$ is the path-loss intercept, d_0 is the intercept distance, β is the path-loss exponent and X_{σ} is a zero mean Gaussian random variable due to shadow fading or slow fading. Path-loss model can be characterized once the slope and intercept of the linear power-law expression are determined [24]–[26].

Similarly, measurement-based path-loss model consists of distance value and corresponding received signal strength indicator (RSSI) value. Both distance (independent) and RSSI (dependent) true values are susceptible to errors due to measurement error and disturbance, respectively. The difference between the observed value and true value in a variable is called measurement error, which can happen in many ways depending on the measurement scenario such as error in GIS map, GPS error in measurement devices, human error and moving targets in vehicular communication. Disturbance in dependent variable is caused by arbitrary randomness such as shadow fading or slow fading. Knowing the source of error in dependent variable and the nature of it is beyond the scope of this chapter. It is shown in the next section that any type of measurement

error or disturbance in dependent variable will add up in one error term; therefore, to find a regression line, there is no need to distinguish source of error. Any type of fading that happens in path-loss modeling can be treated like an error in dependent variable with mean of zero and standard deviation of σ_ε [27].

In propagation and channel modeling, ordinary least square (OLS) is a method, which has been using in the literature to fit a regression line with a RSSI dataset [25], [28]. OLS method works under two assumptions: (I) OLS can determine the slope and intercept of the regression line when there is an error in dependent variable, (II) the distance span of measured data should be long enough to overcome any fluctuation in the log of dependent variable which can conceal the linear relationship of dependent and independent variables [29]. OLS performance is acceptable, when the distance span is in the range of hundreds and thousands of meters which is the case in microcell and macro cell environments and node location accuracy is within few meters. When both independent and dependent variables are subject to error in a short distance span, OLS leads to attenuation bias and inconsistent estimation of slope and intercept [30]–[32].

When both independent and dependent variables are subject to error, error-in-variable (EIV) models are used to find a regression line fitted to data. While EIV models have been used in other fields of study such as chemistry [33], [34] and economics [35], [36], to the best of our knowledge they have never been used previously for propagation channel modeling. Suitable EIV models can be applied to short range propagations like wireless sensor networks and vehicular technologies where distance span may be small and exact position of individual nodes is difficult to obtain [37]–[39]. In short-distance spans with presence of errors in distance, RSSI and logarithmic values of distance get less correlated and slope estimator in OLS get closer to zero whereas EIV is not affected.

EIV models can be classified in two main categories: functional and structural models. When true values of independent variable are unknown, functional EIV models are helpful. Structural models can be applied when true values of independent variables are independently and identically distributed random variables [40], [41]. As measured variables in the path-loss model usually follow a uniform or normal distribution, the structural EIV models can be applied for path-loss modeling. While several methods exist in the literature for structural EIV models, second order method of moments (MoM) EIV method is used in this chapter [40], [42], [43].

This chapter analyzes the performance of both MoM and OLS models for path-loss model estimation in the presence of distance error and shadow fading. This study is implemented over several synthesized path-loss datasets, which are created similar to practical measured datasets of a propagation model. For different distance spans, without loss of generality, slope and intercept values are chosen randomly. Next, to generate distance error, synthesized normal random errors with zero mean and a range of different standard deviations are applied to the distance spans. Finally, OLS and MoM are applied to datasets to estimate slope and intercept.

The objectives of the chapter are:

- To compare the performance of MoM and OLS in estimation of path-loss slope with presence of distance error and in a range of distance span and shadow fading.
- Propose a decision-making metric called “method selector index”, based on the ratio of standard deviation of distance span over distance error in given path-loss dataset; through which one can decide on whether to use MoM or OLS method.
- Investigate the effect of shadow fading on estimation of path-loss slope with MoM.

The rest of the chapter are structured as follows: Section 2.2 details the concepts of error in two dependent and independent variables and explain why MoM method outperforms OLS.

Section 2.3 covers the second order MoM algorithm, its admissibility conditions and algorithm to synthesize datasets. Results of the simulation are presented in Section 2.4. Finally, Section 2.5 concludes the chapter.

2.2 Concept

2.2.1 Regression Line

The general form of a regression line shows a linear relation between true values of dependent and independent variables presented as y and x :

$$y = \alpha + \beta x + e \quad (2)$$

In (2), β and α are the slope and intercept, respectively and e is the error term. The error term e is completely random and happens due to the changes of other independent variables called exogenous variables, which are not considered in the regression model under study. Error term e can be considered uncorrelated with both dependent and independent variables. Let x' and y' denote the measured values of independent and dependent variables, respectively defined as

$$x' = x + \delta \quad (3)$$

$$y' = y + \varepsilon \quad (4)$$

where, both true and measured values are related with δ and ε which are Gaussian random variables with mean of zero and standard deviation of σ_δ and σ_ε , respectively. It is assumed that these measurement errors are not correlated with both dependent and independent variables. In this chapter, while shadow fading is modeled as measurement error in the dependent variable, other types of probable error during measurement are ignored. In case of measurement error, true values are not available and x' and y' are related through the regression line equation as:

$$y' = \alpha + \beta(x' - \delta) + \varepsilon + e \quad (5)$$

2.2.2 Ordinary Least Square Method

The slope and intercept estimator in OLS are defined as $\hat{\beta}$ and $\hat{\alpha}$, respectively as

$$\hat{\beta} = \frac{\text{Cov}[x,y]}{\text{Var}[x]} \quad (6)$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} \quad (7)$$

where \bar{x} and \bar{y} are the mean values of n samples. It should be noted that an acceptable number of samples results in better model estimation and reduction of inference errors.

$$\text{Cov}[x,y] = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \quad (8)$$

$$\text{Var}[x] = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (9)$$

2.2.3 Cases of Error

Three possible cases for error in dependent and independent variables are investigated below:

- Case 1: Error in the dependent variable

In this case $\delta = 0$ and equation (5) reduces to:

$$y' = \alpha + \beta x + \varepsilon + e \quad (10)$$

Let $E = \varepsilon + e$ be the new error term with zero mean and greater standard deviation compared to e and ε . Similar to both e and ε , E is assumed to be uncorrelated with either of the variables. The new error term in the independent variable keep the slope estimator consistent and unbiased but affect the intercept estimator. In the case of path-loss modeling, we can consider both shadow fading/slow fading or any measurement error in RSSI as one error term in independent variable. Since rate of change in RSSI by moving for a unit of distance or simply slope is of great importance compared to intercept, it is acceptable to use OLS model for this case. Increasing the number of the sample will lead to better estimate and decrease the inference errors.

- Case 2: Error in both variables

Considering the case with errors in both variables, which are uncorrelated with both the observed and the measured values.

$$y' = \alpha + \beta x' + (\varepsilon - \beta\delta) \quad (11)$$

Covariance of the new error term with x' as follows,

$$\text{Cov}(x', \varepsilon - \beta\delta) = E[(x + \delta)(\varepsilon - \beta\delta)] = -\beta\sigma_\delta^2 \neq 0 \quad (12)$$

reveals that correlation of the new error term and x' is not equal to zero, unless for $\beta = 0$ or occurrence of no measurement error. Equation 12 shows the classic error in variables problem in which the new error term is correlated with the independent variable. This leads to a biased OLS slope estimator.

$$\text{plim } \beta = \frac{\text{cov}(x', y')}{\text{var}(x')} = \frac{\text{cov}(x', \alpha + \beta x' + (\varepsilon - \beta\delta))}{\text{var}(x')} = \beta \left(1 - \frac{\text{cov}(x', \delta)}{\text{var}(x')} \right) = \beta \left(\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_\delta^2} \right) \quad (13)$$

$\text{plim } \beta$ is called attenuation biased since slope estimation is always lower than true estimate of slope and it will get closer to zero as σ_δ increases. Since in path-loss modeling, measurement of the distance between the transmitter and receiver is prone to error, OLS cannot be a good solution to estimate the slope and intercept of the path-loss model.

- Case 3: Correlated error terms

This scenario happens when the error terms are correlated with the true value and observed value of dependent and independent variables. This implies that all the estimators are biased and OLS should not be applied. This is not the case for measurement-based path-loss modeling and it is out of the scope of this chapter.

2.3 Methodology

In measurement-based path-loss modeling, researchers are dealing with homoscedasticity error term in independent variable and both Maximum Likelihood (ML) and MoM can be applied to

find the estimators. These two methods can be applied interchangeably, though independent variable must have normal distribution in ML whereas it can have other types of distributions in MoM. This is the most important advantage of the MoM over the ML in path-loss modeling due to the possibility of uniform distribution of distance. Also, theoretical asymptotic variance and covariance of the estimators in MoM can be easily calculated for further statistical analysis such as students' t-test and making confidence interval [42], [43].

Based on the second order MoM, five estimating equations are:

$$E[x'] = E[x] = \hat{\mu} \quad (14)$$

$$E[y'] = E[y] = \hat{\alpha} + \hat{\beta}\hat{\mu} \quad (15)$$

$$\text{Var}[x'] = \text{Var}[x] + \text{Var}[\delta] = \hat{\sigma}_x^2 + \hat{\sigma}_\delta^2 \quad (16)$$

$$\text{Var}[y'] = \text{Var}[\hat{\alpha} + \hat{\beta}x] + \text{Var}[\hat{\epsilon}] = \hat{\beta}^2 \hat{\sigma}_x^2 + \hat{\sigma}_\epsilon^2 \quad (17)$$

$$\text{Cov}[x', y'] = \text{Cov}[x, \hat{\alpha} + \hat{\beta}x] = \hat{\beta} \hat{\sigma}_x^2 \quad (18)$$

where first and second order sample moments on the left side of equations are calculated from measurement data and the corresponding population moments with hat sign on the right side of equations are a function of model's parameter. Here, hat sign denotes population estimators. Note that six parameters in five estimating equations in MoM implies that no unique solution can be found. This is the main problem with MoM and is called "identifiability problem". Two solutions to overcome this problem are: 1) to obtain prior knowledge about one of the estimators; once the value of an estimator is determined, other parameters can be estimated. 2) to obtain more estimating equations with third and fourth moments.

The problem with higher order estimating equations is the necessity to collect more data to have accurate estimation of third and fourth moments, which is neither time efficient nor cost effective. Additionally, as third and fourth moments of random variables are a measure of

skewness and kurtosis, respectively, it is important that both dependent and independent variables are adequately skewed and kurtotic [40], [43]. This chapter uses second moments to estimate equations and assumes that prior knowledge of the distance error is available.

Measurement-based propagation modeling might take place with error in distance and RSSI variables. The usual procedure in path-loss modeling is to have uniform distribution for distances and assume zero mean Gaussian normal distribution for distance and RSSI error terms, which are uncorrelated with both independent and dependent variables. In the absence of LOS propagation, corresponding distance of any NLOS propagation deviates from the actual distance between two nodes and causes distance error. Distance error usually happens due to error in GIS map, GPS error, moving targets, human error, etc. and calculating the standard deviation of error when the source of error is determined is a feasible task.

Once σ_δ^2 is known, the slope estimator in MoM can be derived from (16) and (17) as follows:

$$\hat{\beta}_{\text{MoM}} = \frac{\text{Cov}[x',y']}{\text{Var}[x'] - \sigma_\delta^2} \quad (19)$$

This formula is the adjusted version of the OLS slope estimator. The adjustment is the subtraction of distance error variance from the observed distance variance to estimate true value of distance variance.

2.3.1 Admissibility Conditions

To make sure measured-data remains in the range of structural model, the following admissibility conditions which derived from (16), (17), (18) should be satisfied:

$$\text{Var}[x'] > \sigma_\delta^2 \quad (20)$$

$$\text{Var}[y'] > \frac{\text{Cov}[x',y']^2}{\text{Var}[x'] - \sigma_\delta^2} \quad (21)$$

$$\frac{\text{Var}[y']}{\text{Cov}[x',y']} < \hat{\beta}_{\text{EIV}} < \frac{\text{Cov}[x',y']}{\text{Var}[x']} \quad (22)$$

Violating equation (20), (21) can be interpreted as negative variance for true values of distance and RSSI, which in fact violates the assumption in (3) and (4), therefore no EIV method can be used in these circumstances. Equation (22) is the other general admissibility condition of MoM, which shows the range of the estimated value of slope. The upper limit value is OLS of y' on x' and the lower limit value is OLS of x' on y' .

Data sample size can affect the strict MoM admissibility conditions. In the synthesized data set, it is set to 50 when the distance error variance is known a priori [40]. By increasing the number of samples, slope estimator will be less biased and get close to its true value and violation of (22) is less likely to happen. On the other hand, increasing number of samples can prolong the measurement time and increase the cost. It was reported that at a sample size of 50, MoM shows 2% error in slope estimation and less than 20% chance of breaking the last admissibility condition [40].

2.3.2 Synthesized Dataset

To synthesis a dataset that resembles the measurement-based dataset features, four different distance spans with length of 25, 50, 100, and 200 m were chosen. The starting point of all these distance spans were 25 m which is the realistic node distances in short-range wireless technologies. Next, 50 points with a uniform distribution in each distance span were selected. Without loss of generality, a set of values for path-loss slopes as $\beta_{det} = \{-1, -2, -3, -4, -5\}$ were chosen which is called “predetermined slopes” and the corresponding RSSI for all slopes and distance spans values were computed, accordingly. The path-loss intercept was set to zero in this chapter. After synthesizing the true values of RSSI and distance with the known path-loss slope and intercept, two Gaussian random variables with mean of zero and standard deviation of σ_ε and σ_δ were also synthesized and added to true values in order to synthesize the observed values of RSSI and

distance. OLS and MoM were applied to the synthesized dataset. To reduce the impact of randomness in predicting the behavior of slope in the synthesized dataset, the complete synthesizing algorithm was repeated 10 times.

2.4 Results

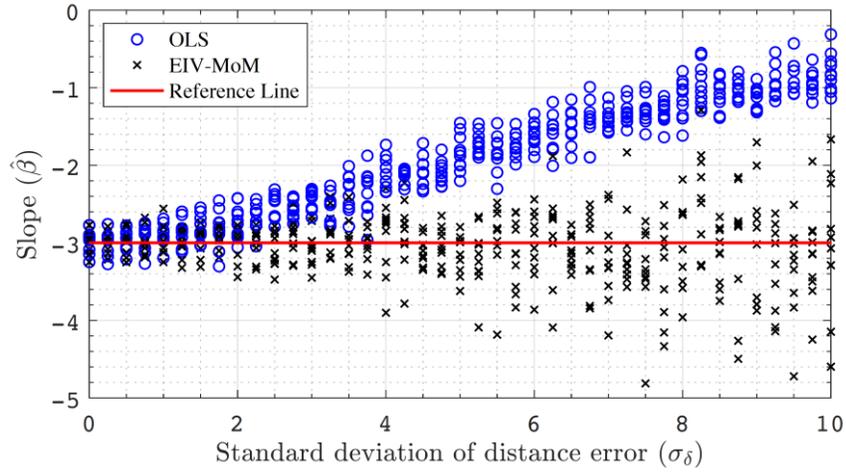
In the first subsection, the preliminary results of the simulation are shown and a description of how slope in path-loss model estimation deviates from true values of slope with OLS is shown. Furthermore, it is shown how MoM outperform OLS and gives accurate results in the presence of distance error and shadow fading.

In the second subsection, a metric for the distance spans of 25, 50 and 100 m that can be used to choose between OLS and MoM in slope estimation for a measurement-based path-loss data is introduced. In the third subsection, the effect of Shadow fading on accuracy of MoM method is introduced.

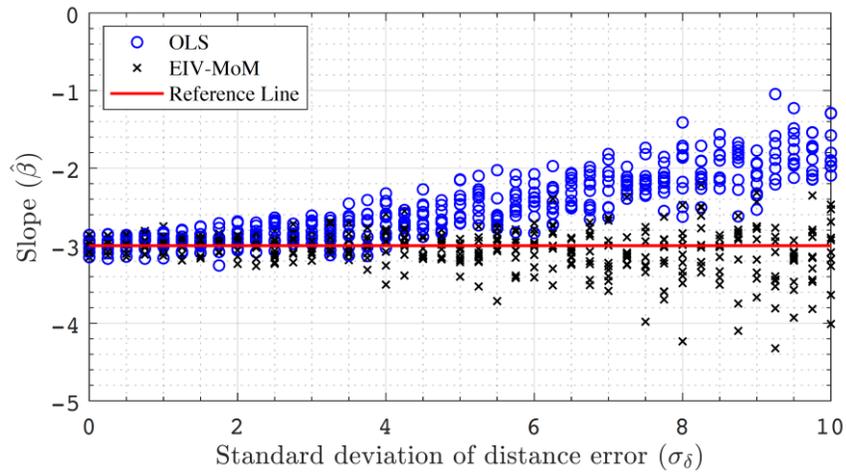
2.4.1 Performance of MoM vs OLS

The preliminary results of estimated slope obtained with OLS and MoM as a function of σ_δ for various distance spans are provided in figure 2.1. To avoid repeating similar results, figure 2.1 only presents the data for $\beta_{det} = \{-3\}$. In this section, $\sigma_\epsilon = 1$ and $n = 50$ is assumed. It can be seen that OLS fails to give an accurate estimate of slope as σ_δ increases. Increasing σ_δ , reduces the correlation of RSSI and distance, therefore OLS slope estimation goes toward zero. The impact of distance error on estimated slope using OLS is reduced by increasing the distance span. Furthermore, it is shown in Figure 2.1 (d) that OLS and MoM give almost the same estimates in the presence of distance error and can be used interchangeably. Additionally, while MoM spreads out around the true values of slope especially for larger values of σ_δ , it spread less for larger distance spans. It is interesting to note that slope estimation of OLS is swaying away around true

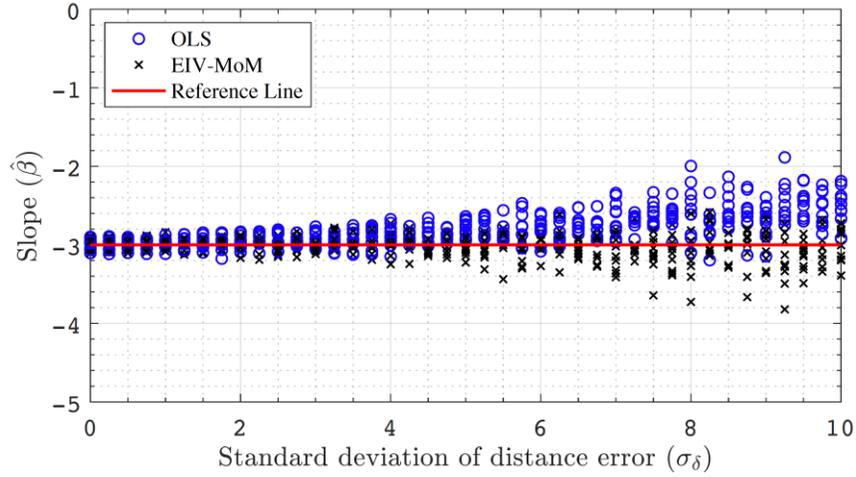
value with a specific pattern as the σ_δ is increasing whereas for MoM it is varying centered on the true value.



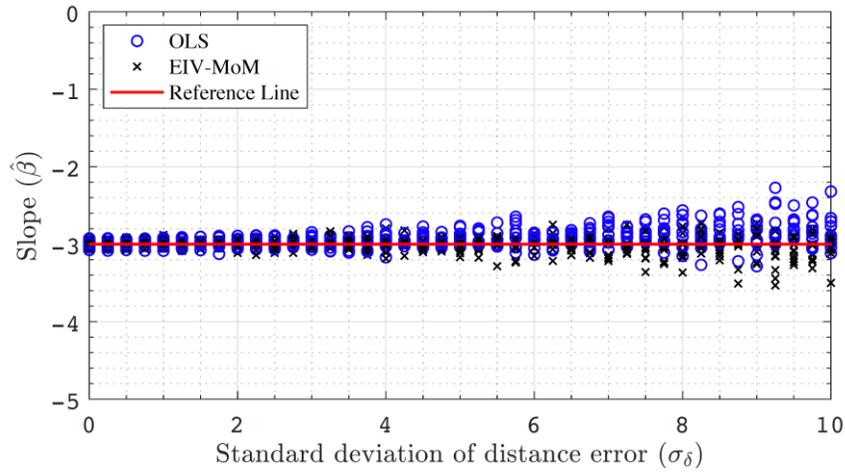
(a)



(b)



(c)

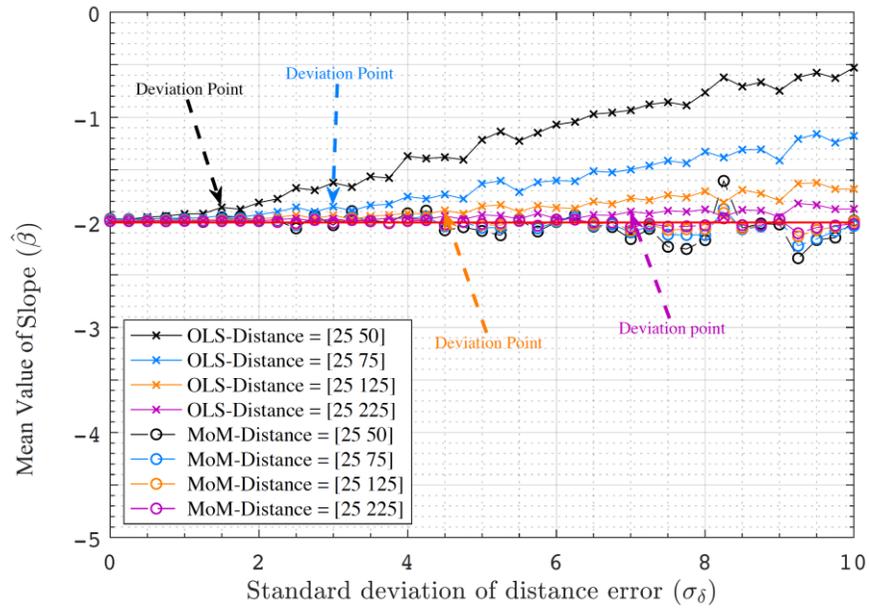


(d)

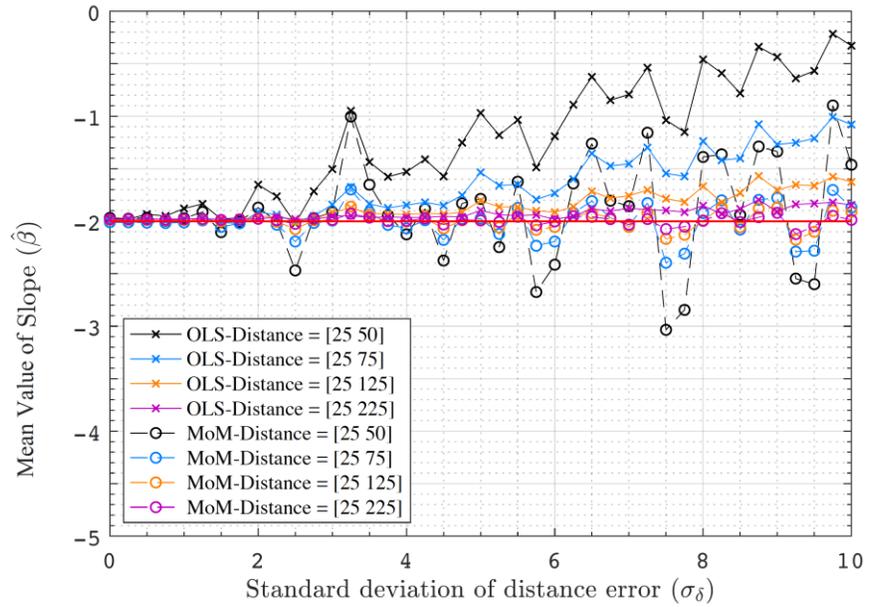
Figure 2.1. $\hat{\beta}$ versus σ_δ for OLS and MoM, (a) Distance span= [25 50] m, (b) Distance span= [25 75] m, (c) Distance span= [25 125] m and (d) Distance span= [25 225] m

Figures 2.2 and 2.3, part (a) and (b), show the mean value of slope estimator both for OLS and MoM taken from 10 samples as a function of σ_δ . It can be seen OLS starts deviating from true value of slope along σ_δ . $\sigma_{\delta_{critical}}$ is defined where the subtraction of estimated slopes using MoM and OLS reach 0.1 and after that point as σ_δ increases, deviation of OLS and MoM increments. As the distance span gets shorter, $\sigma_{\delta_{critical}}$ gets smaller, which causes great amount of slope error in short-range wireless communication. In other words, for short distance spans even slight amount

of distance error can affect the path-loss model.

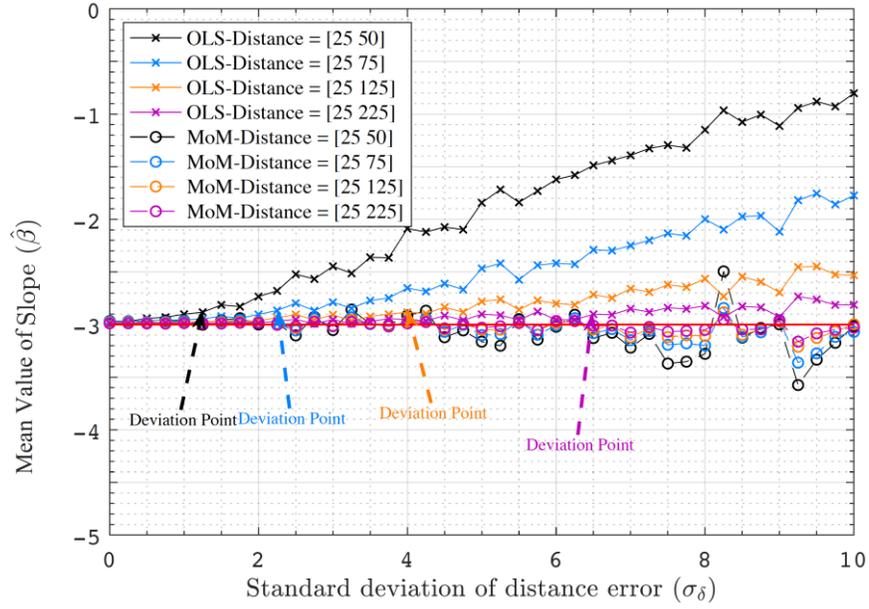


(a)

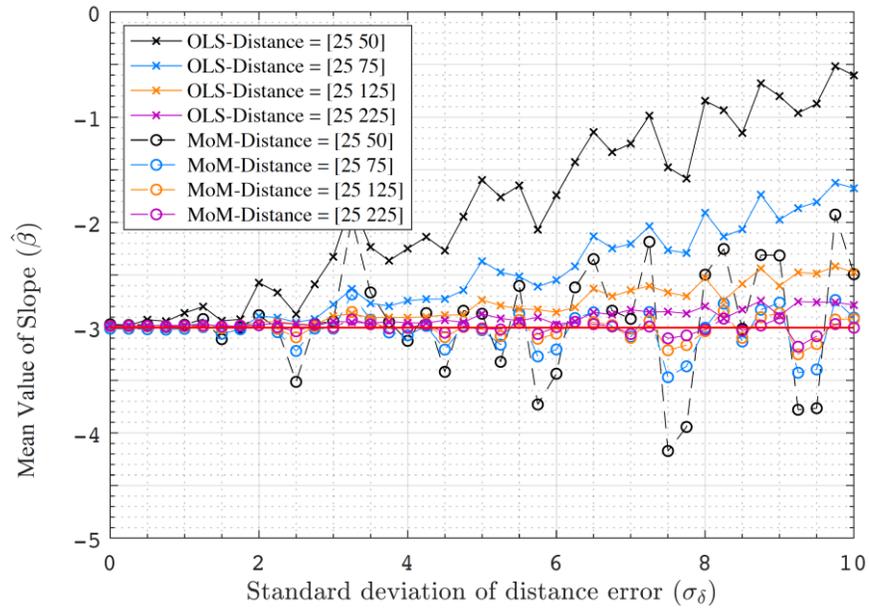


(b)

Figure 2.2. Mean value of slope versus. σ_δ for (a) $\sigma_\epsilon=1$ and $\beta_{\text{det}}=-2$ and (b) $\sigma_\epsilon=9$ and $\beta_{\text{det}}=-2$



(a)



(b)

Figure 2.3. Mean value of slope versus. σ_δ for (a) $\sigma_\varepsilon=1$ and $\beta_{det}=-3$, and (b) $\sigma_\varepsilon=9$ and $\beta_{det}=-3$

The mean value of slope in MoM follows true value of slope and its standard deviation is close to zero for $\beta_{det} = -2, -3$ and for $\sigma_\varepsilon = 1$.

In the case of $\sigma_\varepsilon = 9$, the mean value of slope in MoM still represent the true value of slope but

the standard deviation of slope for short distance spans increases compared to the corresponding value for $\beta_{det} = -2, -3$. Table 2.1 and 2.2 shows the mean value and standard deviation of slope taken over the $\sigma_\delta = [0\ 10]$ both for OLS and MoM.

Table 2.1. Mean and standard deviation of MoM slope over σ_δ for $\beta_{det}=-2$.

MoM\OLS	Span = [25 50]	Span= [25 75]	Span= [25 125]	Span= [25 225]
Mean of Slope ($\sigma_\varepsilon = 1$ and $\beta_{det} = -2$)	-2.02\ -1.28	-2.01\ -1.65	-2.00\ -1.85	-2.00\ -1.94
Standard deviation of slope ($\sigma_\varepsilon = 1$ and $\beta_{det} = -2$)	0.11\ 0.48	0.06\ 0.26	0.04\ 0.10	0.02\ 0.04
Mean of slope ($\sigma_\varepsilon = 9$ and $\beta_{det} = -2$.)	-1.90\ -1.21	-2.00\ -1.65	-1.99\ -1.85	-1.99\ -1.93
Standard deviation of slope ($\sigma_\varepsilon = 9$ and $\beta_{det} = -2$.)	0.47\ 0.56	0.16\ 0.30	0.07\ 0.12	0.03\ 0.05

Table 2.2. Mean and standard deviation of MoM slope over σ_δ for $\beta_{det}=-3$

MoM\OLS	Span = [25 50]	Span = [25 75]	Span = [25 125]	Span = [25 225]
Mean of Slope ($\sigma_\varepsilon = 1$ and $\beta_{det} = -3$)	-3.05\ -1.93	-3.03\ -2.49	-3.01\ -2.79	-3.01\ -2.91
Standard deviation of slope ($\sigma_\varepsilon = 1$ and $\beta_{det} = -3$)	0.17\ 0.72	0.10\ 0.39	0.06\ 0.16	0.04\ 0.06
Mean of slope ($\sigma_\varepsilon = 9$ and $\beta_{det} = -3$.)	-2.93\ -1.87	-3.03\ -2.49	-3.00\ -2.78	-2.99\ -2.90
Standard deviation of slope ($\sigma_\varepsilon = 9$ and $\beta_{det} = -3$.)	0.51 \ 0.79	0.18\ 0.43	0.09\ 0.18	0.05\ 0.07

As σ_ε increases, the MoM is not consistent at the vicinity of the slope's true value and starts oscillating. This oscillation gets boosted when distance span decreases. Eventually, at higher values of σ_δ , OLS slope estimation moves towards zero and MoM oscillation increases around the true value of slope.

Based on the results provided in this section it can be concluded that:

- When $\sigma_\delta < \sigma_{\delta_{critical}}$ for each distance span both MoM and OLS give the same slope estimate and it is more convenient to use OLS instead of MoM.
- For each distance span $\sigma_{\delta_{critical}}$ remains constant by increasing the shadow fading which is further discussed in the next subsection.

- For each distance span when $\sigma_\delta > \sigma_{\delta_{critical}}$, MoM shows better performance in estimating path-loss models.
- For the shortest distance span, which is equal to 25 m, once σ_δ gets bigger than a certain threshold MoM cannot represent the true value of slope anymore due to oscillation and deviation from the true value of slope. This threshold is further quantified in the following subsections.

As distance span increases, OLS and MoM are less susceptible to distance error and they perform almost equally. For any distance span, larger than 100 m OLS can be used even in the presence of distance error.

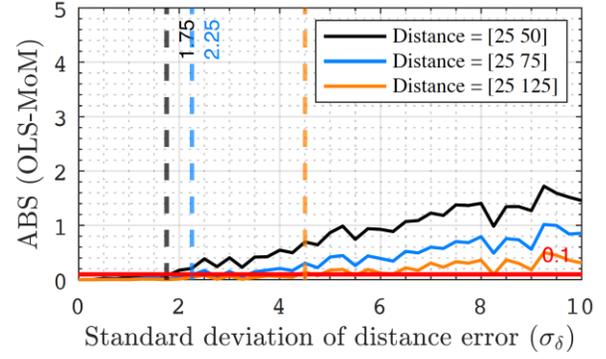
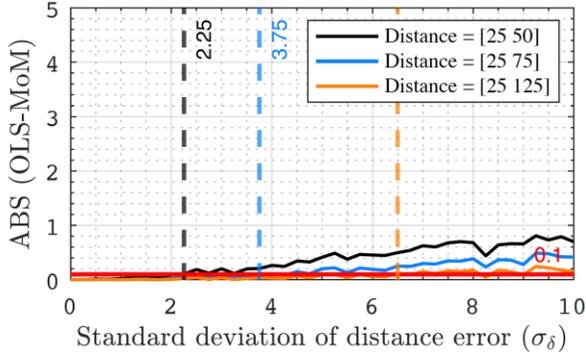
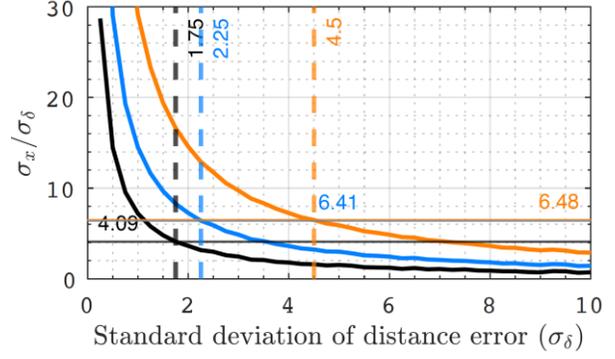
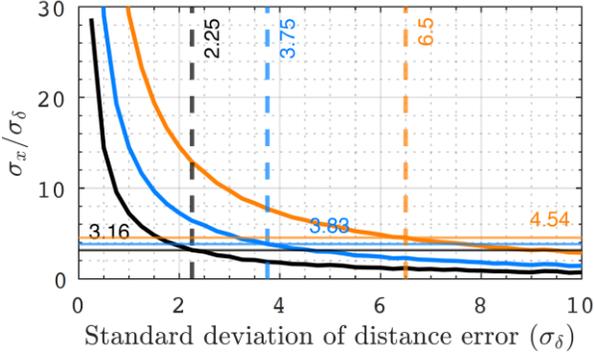
2.4.2 Applicability of MoM

In this section, an empirical metric called “Method Selector index” is computed. Given a measurement-based path-loss data, this metric can be used to decide whether to use MoM or OLS. This metric calculates $\hat{\sigma}_x/\sigma_\delta$ and gives the user a value which shows how real path-loss data stands regarding deviation points. $\hat{\sigma}_x$ is the estimated standard deviation of true values of distance which can be calculated from (16). Once σ_δ is calculated for a given path-loss measurement scenario, method selector index can be determined. Figure 2.4 (a-e), shows the change in the metric at the top and change in subtraction of OLS from MoM at the bottom with regards to σ_δ . In a given path-loss dataset and for a specific distance span, by calculating $\hat{\sigma}_x/\sigma_\delta$, if the value of index stands at the right side of the deviation point, MoM will give accurate estimates of slopes. Otherwise OLS accuracy would be enough and can be used instead.

Figure 2.4(f) is plotted with the method selector values at the $\sigma_{\delta_{critical}}$, with respect to the slopes. The plot in figure 2.4 (f) is for distance spans smaller than 100 m. The maximum value that

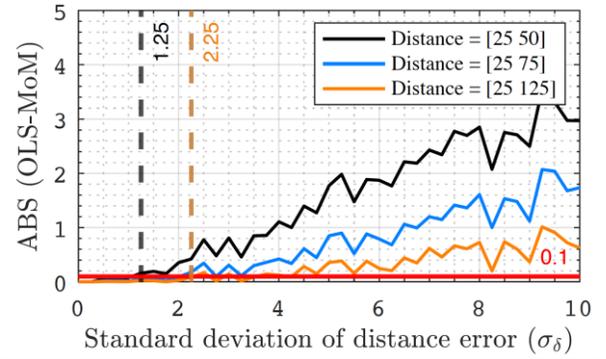
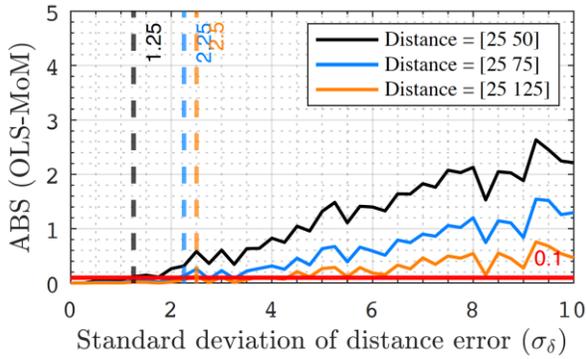
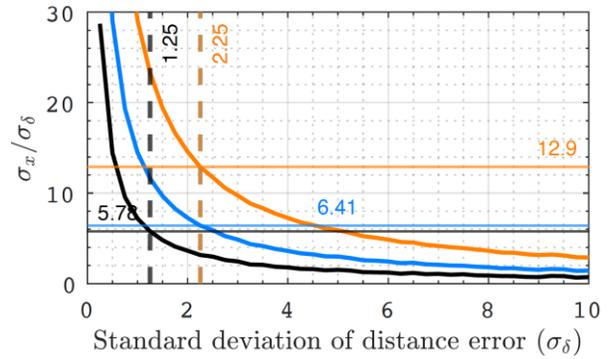
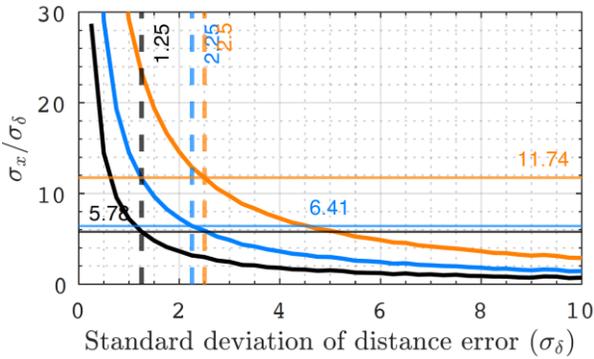
method selector index can get in each distance span is set as a threshold to decide whether to use MoM or OLS. In a given path-loss dataset with a specific distance span, MoM should be used if the obtained value of method selector index is smaller than the maximum value, and if it is greater than the maximum value, OLS should be used.

For the other distance spans with different ranges and standard deviations, the critical value of regression finder can be interpolated. Although the uniform distribution for the distance data was chosen, similar results can be obtained for chi and normal distributions. It is worth mentioning that before calculating the method selector for a measurement-based dataset, one should test the admissibility conditions in equation (20) and (21) for the measurement data.



(a)

(b)



(c)

(d)

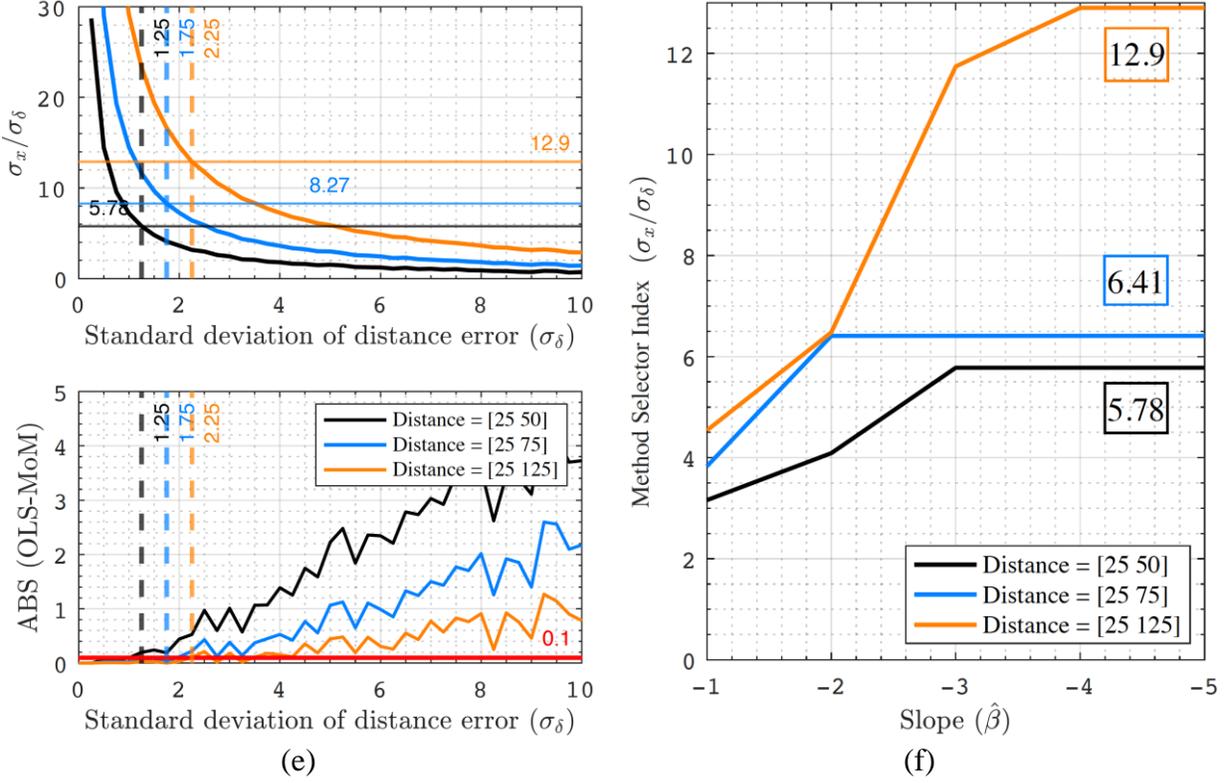


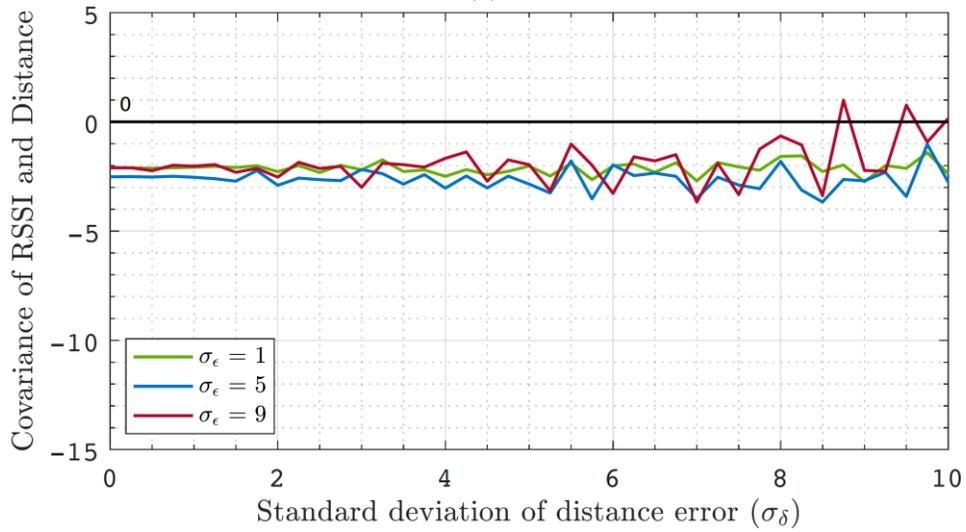
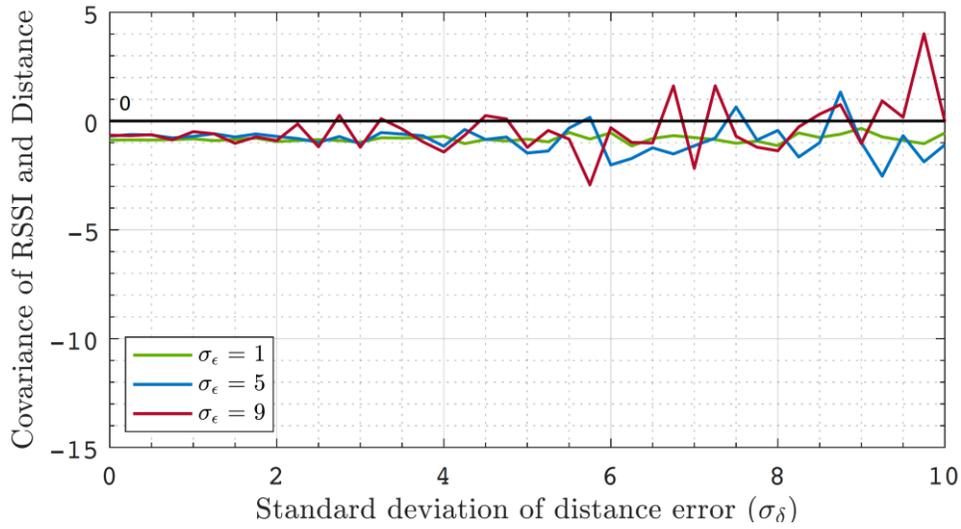
Figure 2.4. (a) $\beta_{\text{det}}=-1$ (b) $\beta_{\text{det}}=-2$ (c) $\beta_{\text{det}}=-3$ (d) $\beta_{\text{det}}=-4$ (e) $\beta_{\text{det}}=-5$ (f) threshold of method selector index versus slope

2.4.3 Effect of Shadow Fading on MoM

In this section, the effect of shadow fading on the performance of MoM is investigated. Figure. 2.5 and 2.6 show the covariance of RSSI and distance versus σ_δ for the synthesized dataset and for distance spans of 25 and 50 meters for three different values of shadow fading. Increasing distance error and shadow fading in short range wireless communication makes the path-loss data less correlated; therefore, the covariance of RSSI and distance goes toward zero. Also increasing σ_δ in the presence of great shadow fading leads to oscillation of the covariance.

As the slope in (19) is calculated with the covariance of the RSSI and distance, here the effect of shadow fading on the covariance is investigated. Two hard and soft thresholds based on the covariance of RSSI and distance is determined. Covariance of 0 is set for the hard threshold beyond

which MoM cannot estimate the true value of slope. A positive value for covariance of RSSI and distance means that combination of shadow fading and distance completely change the behavior of path-loss data. Covariance of -1 is set for the soft threshold. Once the soft threshold is violated there is no solid evidence to show whether true value of slope is estimated by MoM or not.



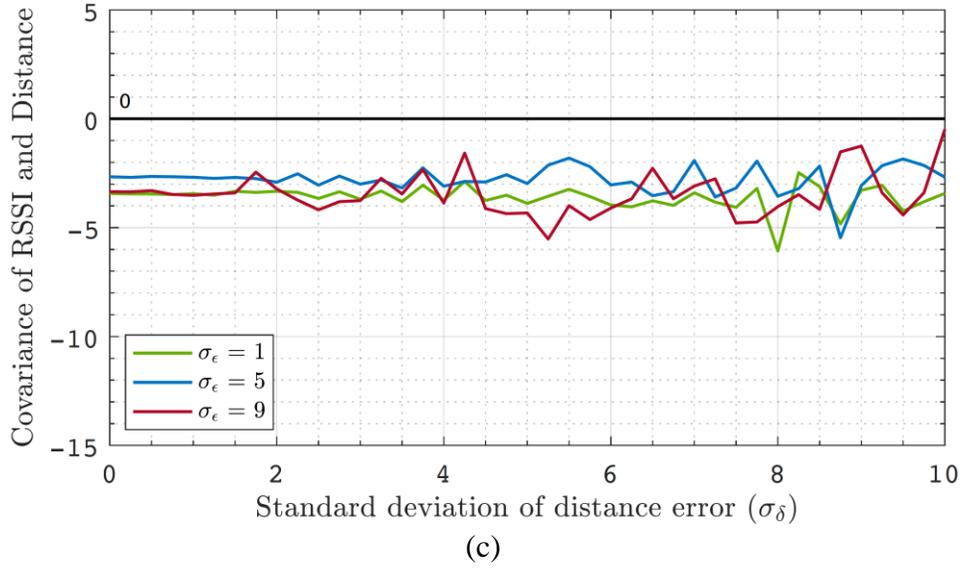
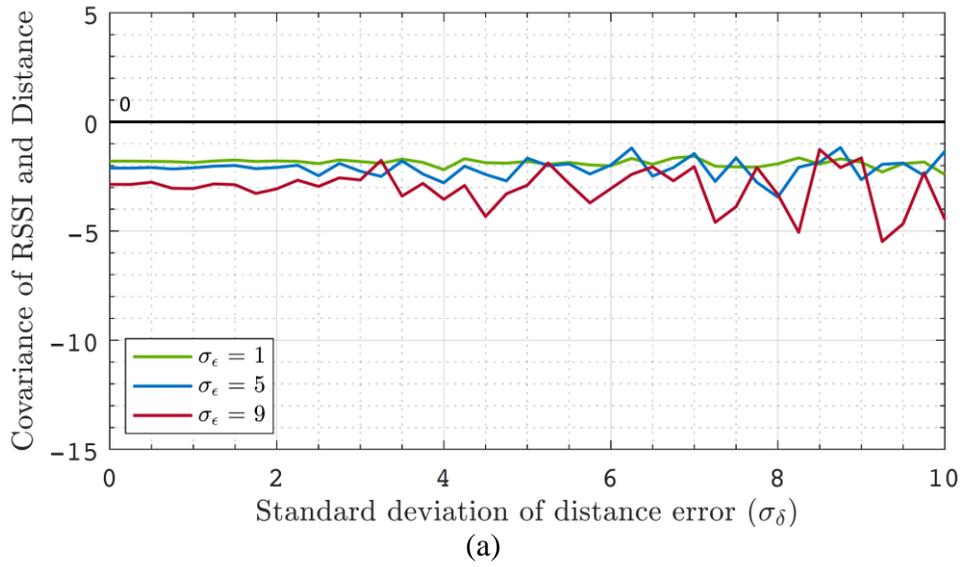
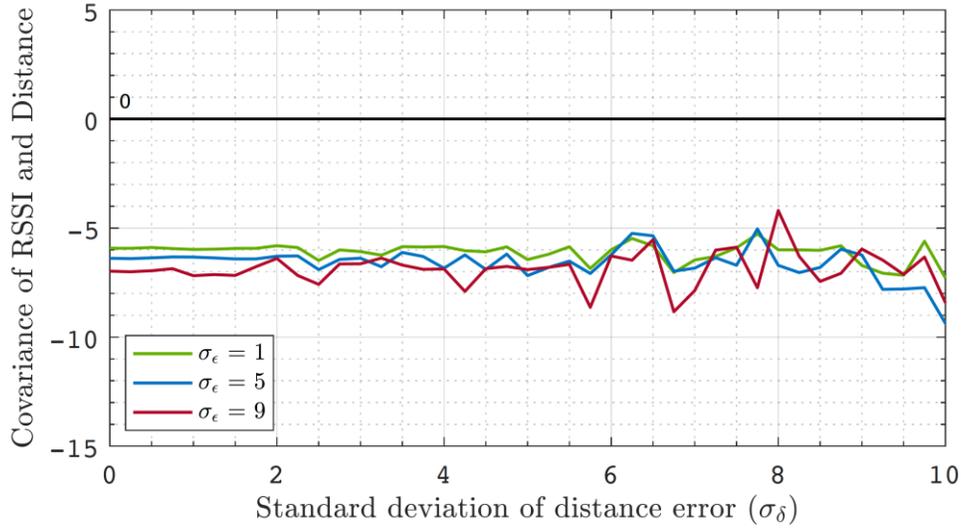
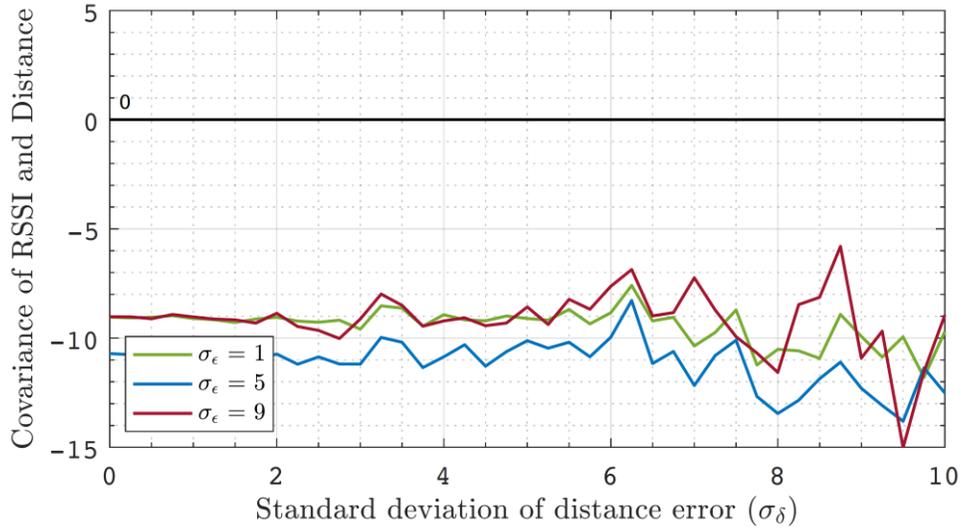


Figure 2.5. Distance span = [25 50] m. (a) $\beta_{\text{det}}=-1$ (b) $\beta_{\text{det}}=-3$ (c) $\beta_{\text{det}}=-5$





(b)



(c)

Figure 2.6. Distance Span = [25 75] m. (a) $\beta_{det}=-1$ (b) $\beta_{det}=-3$ (c) $\beta_{det}=-5$

2.5 Discussion

In this chapter, the performance of MoM and OLS with presence of error in distance was compared with the help of a synthesized dataset. It was shown that for short-range wireless propagation below a distance span of 100 meters in the presence of distance errors and shadow

fading, MoM outperforms the OLS in slope estimation for path-loss modeling. It is also shown how OLS deviates from true values of slope and goes toward zero as distance error increases whereas MoM remains consistent around true values of slope. The chapter has defined deviation point as the point where MoM starts outperforming the OLS. For each distance span, the corresponding value of distance error for deviation point is defined as the critical distance error.

The chapter has also shown the impact of shadow fading on MoM. Based on the results, it can be seen that shadow fading increases the oscillation around true values of slope for MoM but doesn't affect the location of deviation point in each distance span. To show applicability of MoM, the authors came up with a decision-making metric called "Method Selector index", which shows where the data stands regarding to the deviation point. The values of this metric working based on the critical value of distance error give the user a guidance on whether to use MoM or OLS for a real path-loss dataset.

Finally, this chapter obtained the conditions that shadow fading obscures the covariance of RSSI, distance, and correspondingly estimation of path-loss. Increase in the standard deviation of shadow fading and distance error, in addition to the decrease of distance span, weakens the correlation between distance and RSSI values in the measured dataset. The authors came up with hard and soft thresholds for the covariance of RSSI and distance that can be tested in real path-loss dataset. Beyond hard threshold the effect of shadow fading is high enough that MoM cannot estimate the true value of slope. Beyond the soft threshold, there is no solid evidence to show whether true value of slope is estimated or not.

Chapter 3: Long-Term Temporal and Spatial Statistical Assessment of Short-Range Power-Law Path-loss Model in Sub-Urban Environments with BC Hydro Utility Network Dataset.

3.1 Introduction

Path-loss models are essential in estimating link reliability and mutual interference in wireless mesh networks (WMNs). Path-loss is a metric to measure the performance and connectivity of communication link in WMNs. A comprehensive and accurate path-loss model which shows the variability of communication channel with time, location, and distance can be used by network operators and manufacturers to analyze multi hop communication in sensor and mesh networks. Such models can be incorporated in tuning and optimizing consumed powers by each node in a WMN.

Early path-loss models were concerned with point to multi-point UHF and VHF land mobile services [44]. More recently, wireless mesh network path-loss model has been the focus of research due to development of device to device (D2D) and device to infrastructure (D2I) communication links [24], [25], [37]–[39], [45], [46]. The majority of these path-loss models cannot be accurate representatives for modeling mesh hopping network or a WMN. Fundamental discrepancies between communication channels in literature and WMNs such as BC Hydro’s multi-service grid network (MSGN) are listed below:

- Most of the antennas in smart utility networks (SUN) are low height antenna in confined spaces whereas previous path-loss models have been developed in open spaces with variable antenna height [37]–[39].
- The majority of communication links in SUN are NLOS whereas LOS communication links have been used dominantly in path-loss models [47].

- Path length in SUN is shorter than the one used in previous path-loss models [26].
- Frequency of operation in SUN network follows 802.15.4g IEEE standard whereas some of the path-loss models are established for different range of frequencies [23], [45], [48].
- The models are based on small measurement dataset over a short period of time.

A SUN can be in the scale of a city, province, or country. Therefore, terrain morphology, vegetation density, and urban architecture varies within different segments of network. This can be interpreted as spatial variation of path-loss model within different segments of SUN. Additionally, communication channels which are located in a segment of SUN, vary with time. This long-term temporal variation can be in the order of hours, days, and years. To the best of our knowledge, no existing path-loss model incorporates simultaneous long-term temporal and spatial variability of communication links in path-loss models.

This chapter proposes a comprehensive path-loss model which incorporates the long-term temporal and spatial variability of communication links and is based on the BC Hydro MSGN performance dataset. We use network measured data for three regions with different terrain morphology and vegetation density over the course of three months. In the BC Hydro MSGN, precision error in calculating communication links' path length and network biased toward the best path with higher RSSI values are the two main problems of the collected dataset. Considering the vast number of communication links and number of measured RSSI for each link, these limitations can be solved by filtering the erroneous data. Although network measured data compared to the data measured with lab grade instruments have some constraints, it is the only cost-effective way to capture the spatial and long-term temporal variation of communication links.

The objectives of this chapter are:

- To perform statistical assessment over BC Hydro MSGN data to remove erroneous data, characterize the existing fading in wireless links and determine links' path-loss distributions in time.
- To show a new power law path-loss model incorporates both spatial and long-term temporal variability.
- To estimate the parameters of this new power-law path-loss model based on data collected in three distinct environments.

The remainder of this chapter is organized as follows: Section 3.2 explains how long-term temporal and spatial variability of communication links can be modeled by power-law path-loss model. Section 3.3 demonstrates our methodology to collect and reduce network measured data and our techniques to overcome the challenges of network measured dataset. In Section 3.4, the results of statistical assessment of BC Hydro dataset and our proposed power-law path-loss model are explained. Finally, Section 3.5 concludes the chapter.

3.2 Concept

Log-distance power-law path-loss model is defined as follows [49]:

$$PL(d)=PL(d_0)+10\beta \log_{10} \frac{d}{d_0}+X_\sigma \quad (1)$$

where $PL(d)$ is the path-loss at a distance d expressed in dB, $PL(d_0)$ is the path-loss intercept, d_0 is the intercept distance, β is the path-loss exponent, and X_σ is shadow fading or slow fading error. Path-loss model can be characterized by determining the slope, intercept, and finally shadow fading standard deviation σ of the linear power-law expression.

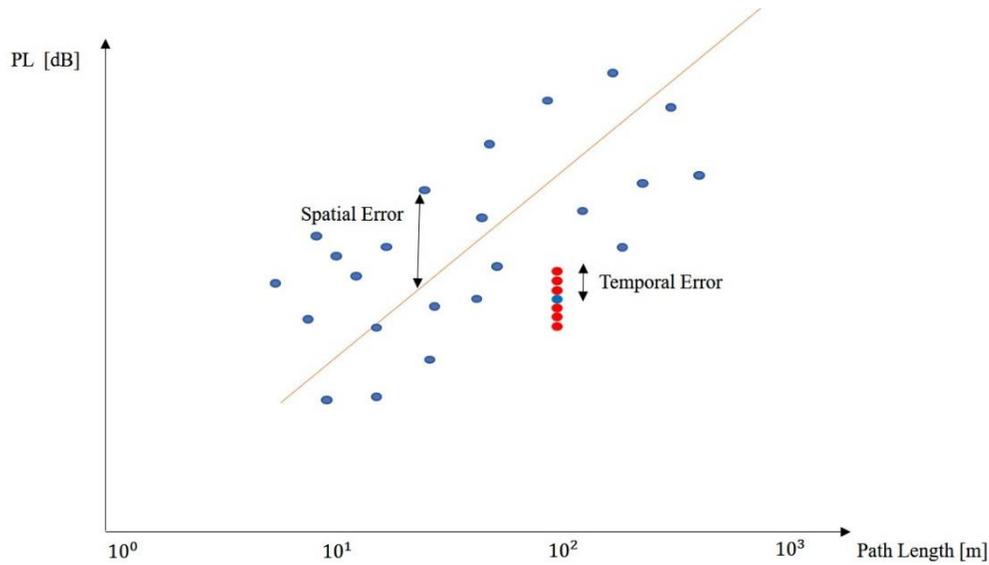


Figure 3.1. Typical Path-loss Model

Figure 3.1 shows typical power-law path-loss model. Blue nodes show RSSI (correspondingly path-loss values by considering the transmitted power and gain of the antennas) and path length values in y and x axis, respectively. Path-loss values are measured by lab grade instruments or the nodes in a mesh network. Red line is the regression line path-loss model fitted with the path-loss data. By analyzing the residual values, it can be seen the shadow fading error (in y axis) follows a normal distribution in logarithmic scale with the mean of zero and standard deviation of σ . This spatial error is due to shadow fading in the environment and varies from one location to another. Capturing shadow fading requires the path-loss values of multiple communication channel with multiple path length values.

In a wireless mesh network, the nodes are stationary, but the environment is changing over time which results in variable path-loss values of some communication links with time. The network operator needs to know long-term temporal variation of communication links to make sure network does not get segmented over time. Calculating the path-loss values (or signal amplitude) of individual communication links over time in linear scale would result in a lognormal

distribution which is the result of large-scale slow fading (shadow fading). In the time varying communication links, this long-term temporal error might increase or decrease the shadow fading error in the y axis. Given the path-loss values of communication links are captured over time consistently, the new path-loss model is defined as:

$$PL(d)=PL(d_0)+10\beta \log_{10} \frac{d}{d_0}+X_{\sigma_s} + X_{\sigma_t} \quad (2)$$

The above formula is valid when β , $PL(d_0)$ and σ_s are calculated with the average path-loss values of each communication links over time. X_{σ_s} is a Gaussian random variable with mean of zero and standard deviation of σ_s . Spatial standard deviation σ_s corresponds to spatial variation of environment in all of the communication links. To analyze the long-term temporal variation of each communication link in logarithmic scale, X_{σ_t} is a normal distribution which is introduced as the long-term temporal distribution of the path-loss model. As path-loss model calculated with path-loss mean values over time, the mean of X_{σ_t} equals zero. σ_t is a random variable which represents the long-term temporal standard deviation of each communication link and has a lognormal distribution. The path-loss model in (2) is based on the assumption of large-scale fading (shadow fading) which happens due to shadowing and obstacle blocking in wireless links.

3.3 Methodology

This D2D dataset is collected from July 20th to October 18th, 2019 over three regions of Richmond, Kamloops, and Terrace in British Columbia. These regions are chosen due to their unique characteristics in their vegetation density, urban architecture, and terrain morphology. Richmond is a region with flat terrain, light foliage, and a mix of high rise, low rise, residential, and farmland. Kamloops is the large community with hilly terrain, light vegetation consistent with its semi-arid climate, and low rise and rancher style dwellings as well low-density

anchorage. The Terrace dataset which includes the City of Terrace and nearby communities of Lakeelse Lake, Thornhill, and Kitselas offers a hilly area with high foliage surrounded by high mountains and temperate rainforest. The following steps are taken to prepare raw data of BC Hydro SUN for further assessment:

- Collecting network RSSI values of links from BC Hydro Network management System (BC-NMS)
- Identifying two node ID in each link
- Downloading location of Node ID from the GIS
- Calculating path length for all links
- Creating three different datasets for all the available RSSI data in each city
- Sorting datasets based on the measurement date of RSSI for each link

After preparing this preliminary dataset, further filtering steps are explained in the next subsection to overcome the challenges of network dataset for path-loss modeling.

3.3.1 Data Filtering

In the BC Hydro MSGN, the locations of smart meters are assumed in the center of the property lot of residential areas in GIS map. This might create a 0 to 5 m error in the path length of links as the nodes are not necessarily located at the center of lots. To decrease path length error, any link shorter than 35 m is removed so the maximum error in shortest path length would be less than 15%.

Additionally, communication links in the dataset are biased toward the “best path” selected by routing algorithm in network layer. “Best path” are defined as communication channels with measured RSSI floor of -95 dBm. The measured RSSI by meter nodes are captured by embedded

SoC device with designed RSSI sensitivity of -120 dBm. Therefore, routing algorithms do not use the links with measured RSSI values below -95dBm. If routing algorithms do not use a link, the RSSI data of that link will not get stored in BC-NMS. If RSSI value of a link is not available for the short period of time within three months, that link is not used for path-loss modeling and is filtered out of the dataset. At this point, the dataset for each city includes links with daily RSSI for three months. As links with time averaged RSSI less than -95dBm are not available in the filtered dataset, the slope and shadow fading in our path-loss model might be underestimated.

For measurement-based path-loss modeling, path lengths are usually chosen in a way to have a uniform distribution. In the case of network dataset, path lengths are fixed and have a normal distribution. To make path length distribution of network dataset closer to uniform distribution, links with repetitive path length values are eliminated from the network dataset.

To understand the scale of BC Hydro smart meter networks, figure 3.2 shows the communication links and nodes in Terrace to show the scale of the collected dataset.

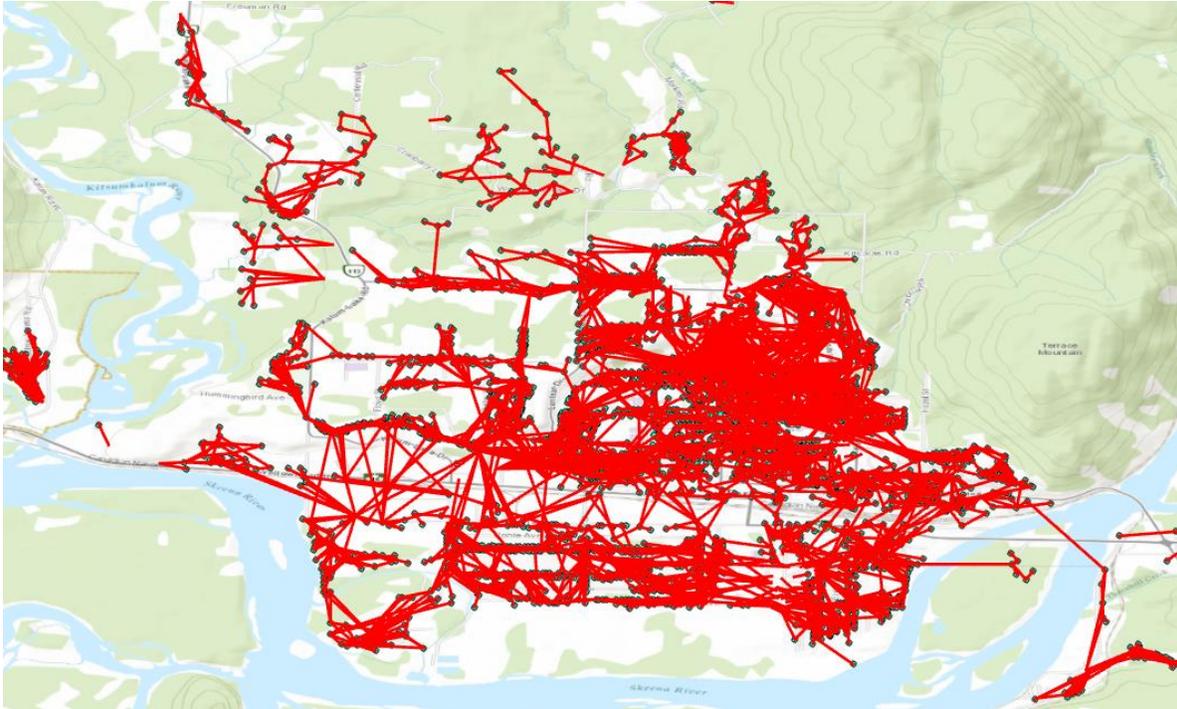


Figure 3.2. Communication links under 1 km and smart meter nodes in the region of Terrace

Table 3.1 shows the scope and scale of communication links and nodes in each region before applying any filter. In the next subsection we characterize the communication channels fading.

Table 3.1. No. of communication links and nodes before filtering

	Richmond	Kamloops	Terrace
No. of Communication links	87,930	44,428	7358
No. of nodes	931,460	585,333	75,625

3.3.2 Fading Characteristics

Communication links in BC Hydro MSGN network follow IEEE 802.15.4g standard for physical and MAC layer. The details characteristics of communication channel can be found in Table 3.2.

Table 3.2. Communication link information

IEEE standard	802.15.4g
Frequency band	902-928 MHz
Communication Channels	64
Channel BW	400 kHz
Modulation	2-FSK
Modulation Index	0.5
Data rate	150 kbps
Occupied BW	225 kHz
Symbol Period	4.4 μ s
Radio receiving time	400 ms to 20 s
Transmitted Power	30 dBm

The RSSI data analyzed in this chapter is captured randomly in a fixed period of 4 hours from 8 pm to 12 am for three months. It is essential to understand the nature of the fading to analyze the dataset. Frequency selective fading cannot happen since the occupied BW is less than the coherence BW, i.e., $1 / \text{RMS delay spread}$. To understand whether slow fading or fast fading happen in the network, distribution of RSSI values for each individual links in linear scale is analyzed and lognormal, Rician, and Rayleigh distributions fitted to the data. We realized lognormal distribution fits the data for the majority of the links and therefore we are dealing with the large scale, slow fading (shadow fading) which happens due to shadowing and blockage in terrain configuration and vegetation foliage [50]. Additionally, the Anderson-Darling goodness of fit was applied to the dataset to verify the lognormality of link's RSSI values in linear scale.

3.3.3 Path-loss Estimation Method

To determine whether OLS method or MoM is suitable to estimate path-loss parameters, the method selector index which introduced in Chapter 2, has been applied to BC Hydro dataset. Given the standard deviation of distance span and distance error, the method selector index can calculate which method gives accurate estimates of parameters. Based on the BC Hydro distance span and distance error OLS would provide accurate estimates for finding the path-loss parameters.

3.4 Results

The results of this chapter are presented in three subsections. In the first subsection, statistical analysis including scale and range of the network dataset is given. In the second subsection, daily values of path-loss slope and shadow fading are presented for three regions to indicate the long-term temporal and spatial variation of path-loss models. Last subsection shows how long-term temporal and spatial variation can be distinguished as two separate term in standard deviation of shadow fading.

3.4.1 Statistical Assessment on the BC Hydro MSGN Dataset

A statistical assessment on the BC Hydro MSGN filtered dataset is provided for three different regions of Richmond, Kamloops, and Terrace. Table 3.3 shows the number of communication links and correspondingly number of smart meter nodes covered in this dataset.

Table 3.3. Number of communication links

	Richmond	Kamloops	Terrace
No. of communication links	101	102	182
No. of nodes	202	204	364

Tables 3.4 and 3.5 show the mean, standard deviation, median, and interquartile range of the path length and RSSI of communication links, respectively for three regions. Figures 3.3 and 3.4 show the distribution of RSSI values and CDF of path length for the communication links in three regions. Due to the applied path length filtering to the datasets, it is clear from figure 3.4 that CDF of path lengths in all regions follows an approximate uniform distribution.

Table 3.4. Communication links path length statistics for each region

Path length\Region	Richmond	Kamloops	Terrace
Min [m]	30.92	52.35	30.89
Max [m]	314.50	348.07	349.65
Mean [m]	117.65	189.61	137.72
Standard deviation [m]	68.17	74.41	68.30
Median [m]	100.02	176.41	121.92
Interquartile range [m]	76.29	114.68	81.46

Table 3.5. Communication links RSSI statistics for each region

RSSI\Region	Richmond	Kamloops	Terrace
Min [dBm]	-109	-108	-102
Max [dBm]	-41	-49	-49
Mean [dBm]	-82.55	-82.73	-79.92
Standard deviation [dB]	10.26	8.47	9.77
Median [dBm]	-84	-83	-81
Interquartile range [dB]	14	11	12

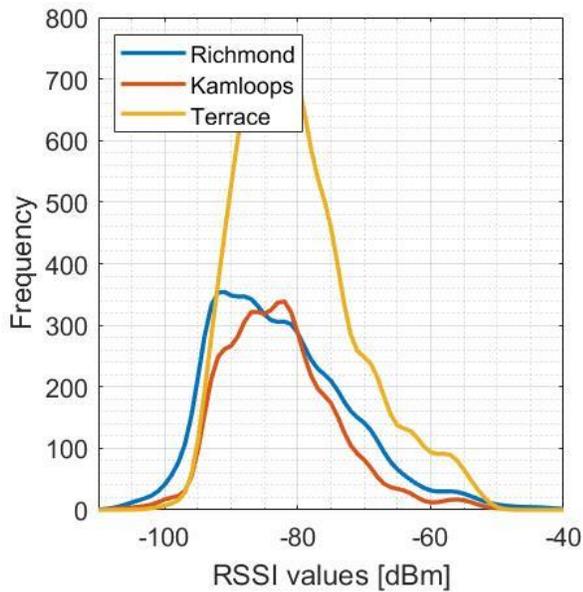


Figure 3.3. Distribution of RSSI data for each city

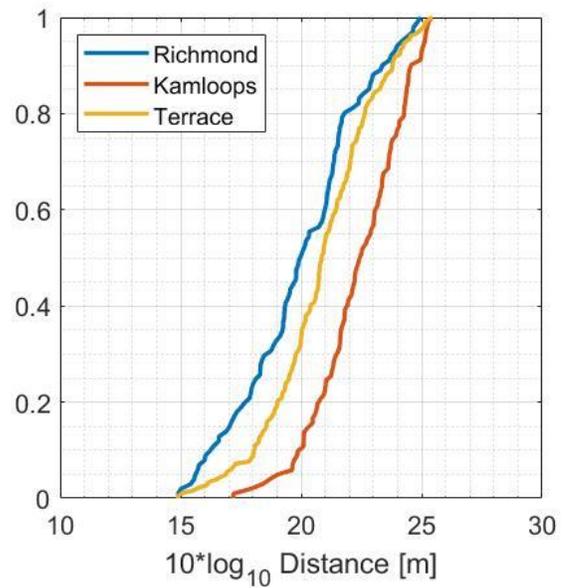


Figure 3.4. CDF of path length

To analyze the lognormality of RSSI values for each individual link, the Anderson-Darling goodness of fit test is applied to daily RSSI values of each individual links at $\alpha = 0.001$. The null and alternative hypothesis include whether linear RSSI values follow the lognormal distribution. In regions of Richmond, Kamloops, and Terrace approximately, 67%, 70% and 76% of the wireless links follows a lognormal distribution. Additionally, Anderson-Darling goodness of fit test at the same significance level is applied to RSSI values of individual wireless links in logarithmic scale to check whether they follow the normal distribution. Same results were observed in each region. The Autocorrelation of path-loss over days in individual links shows random behavior and no correlation is observed between adjacent and near adjacent path-loss data.

3.4.2 Long-Term Temporal and Spatial Variability in Path-loss Parameters

In this subsection, the long-term temporal and spatial variability of communication links based on slope and shadow fading are shown in Figures 3.5 and 3.6, respectively, in three cities. In each city, every slope and shadow fading values are calculated based on the daily values of path-loss

derived from (1) for all links. The range of these values indicate long-term temporal variation in communication links and correspondingly in path-loss modeling. Besides, these values vary from one city to another which indicates the spatial variation in path-loss models. Difference among shadow fading of three regions could be due to unique characteristics of them.

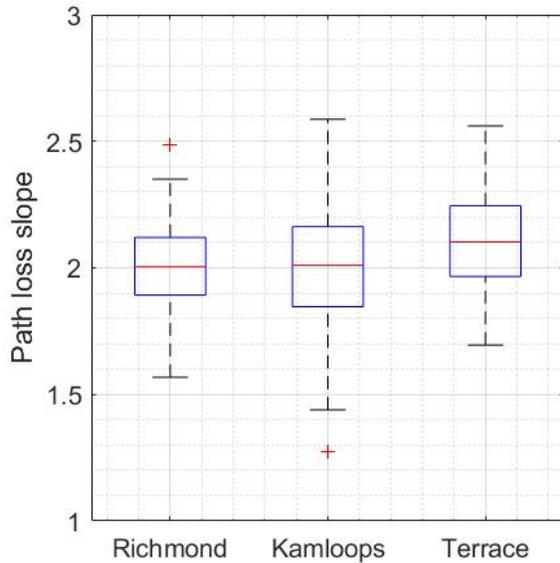


Figure 3.5. Daily value of slope

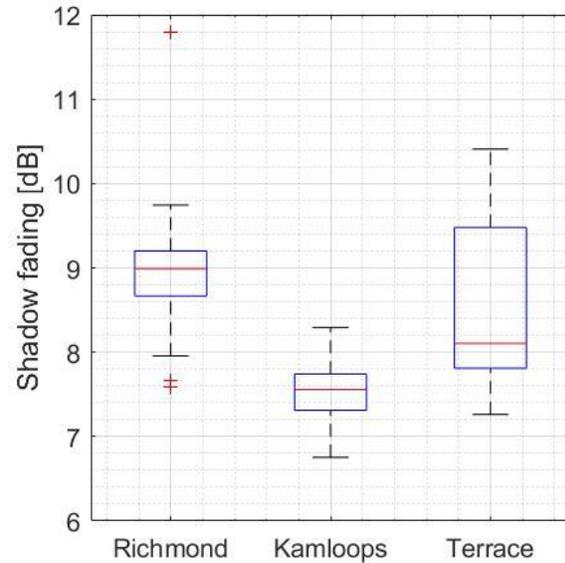


Figure 3.6. Daily value of shadow fading

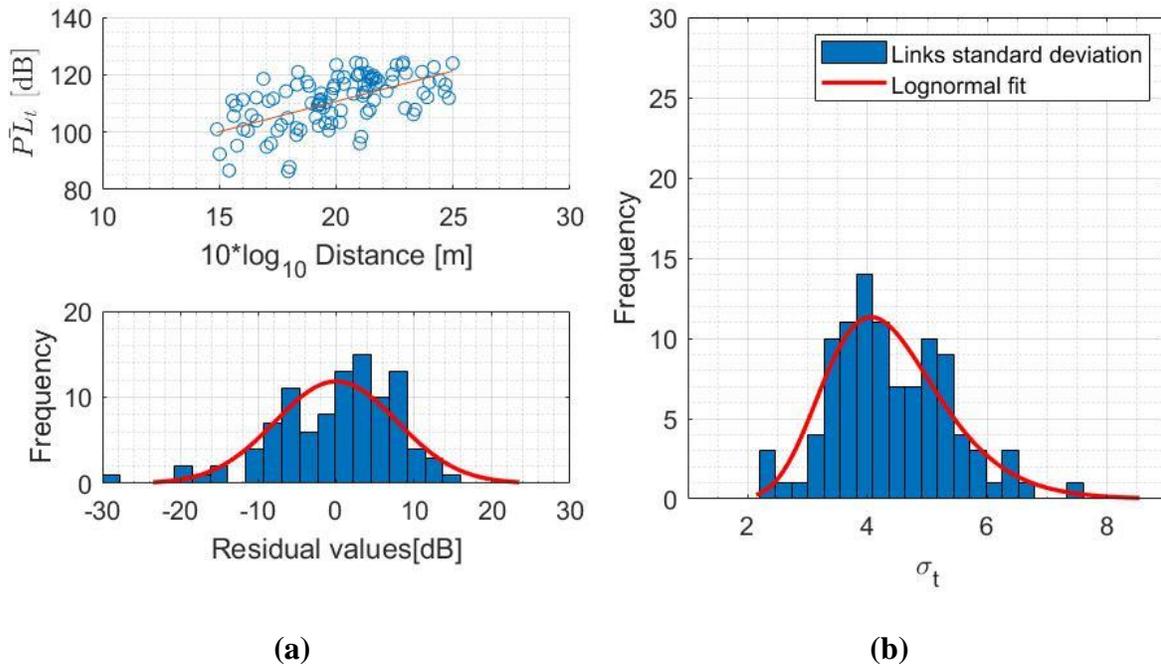
3.4.3 Estimate the Parameters of Path-Loss Model

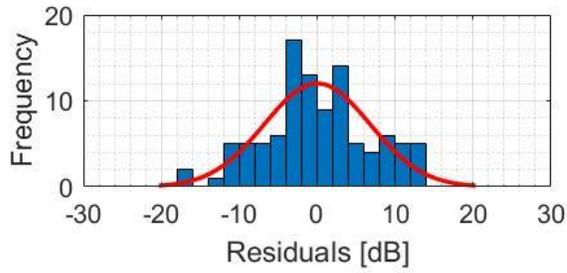
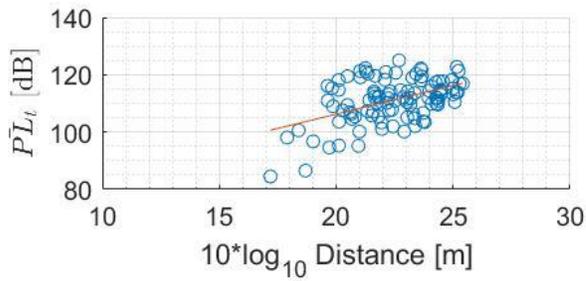
Here our proposed path-loss model based on (2) is explained. Figure 3.7 (a), (c) and (e) show path-loss model, slope, and intercept in all regions. These values are calculated by averaging path-loss values of each link over three months. Additionally, histogram of the residual values is shown for each region. Standard deviation of residual histogram represents the spatial term of shadow fading which is due to the change in the environment of all links in each region.

Long-term temporal variation of path-loss values over time in individual communication links is modeled by their standard deviation. This long-term temporal variation varies from one

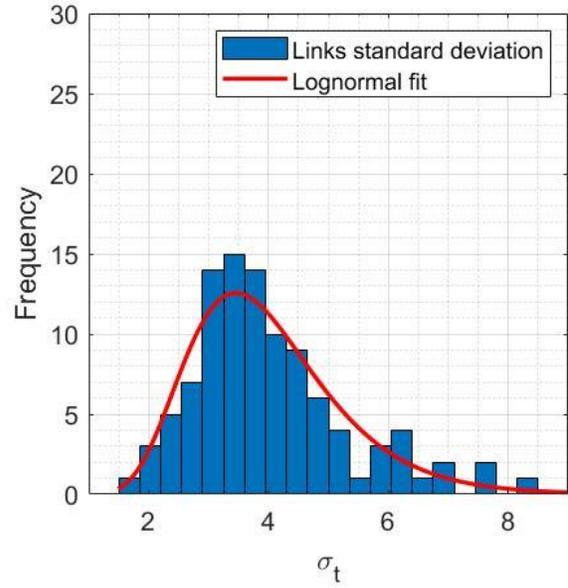
communications link to another. Figure 3.7 (b), (d) and (f) illustrate the distribution of long-term temporal standard deviation for all of the communication links. σ_t here is lognormal distribution which can be characterized by two parameters of location μ and scale σ . Anderson-Darling test is applied to the links distribution of long-term temporal standard deviations at $\alpha = 0.05$ to measure the goodness of lognormal fit. Null and alternative hypothesis include whether the distribution follows the lognormal distribution.

Fading is modeled as two normal random variables with means of zero and standard deviations of σ_s and σ_t . Table 3.6 presents the slope, intercept, spatial fading, standard deviation, long-term temporal fading standard deviation of communication links, and the result of Anderson-Darling goodness of fit test.

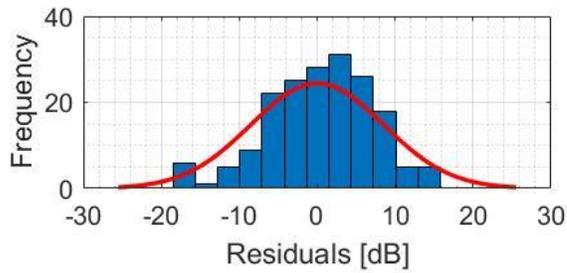
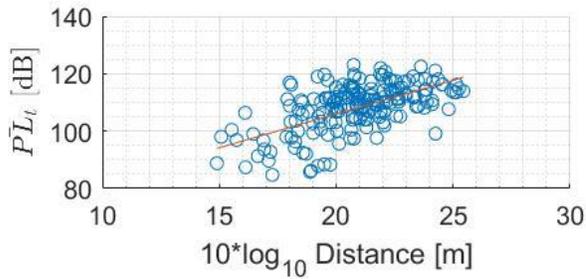




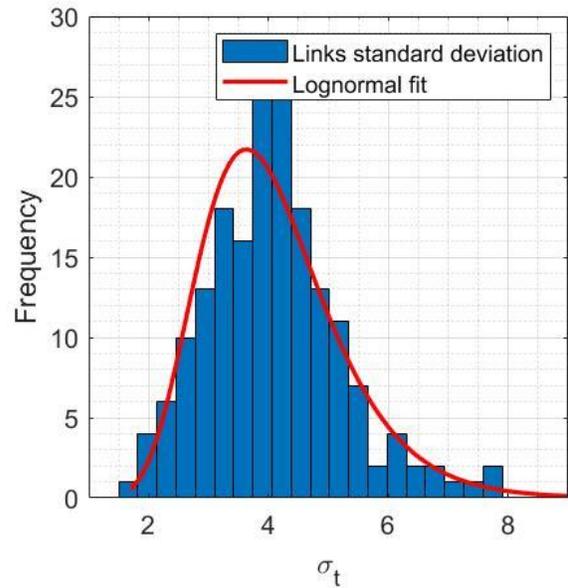
(c)



(d)



(e)



(f)

Figure 3.7. Path-loss model and residual histogram of (a) Richmond, (c) Kamloops, (e) Terrace, Distribution of long-term temporal standard deviation of (b) Richmond, (d) Kamloops, (f) Terrace.

Table 3.6. Path-loss parameters

Regions	Richmond	Kamloops	Terrace
β	2.11	1.99	2.35
$PL(d_0)$ [dBm]	68.44	66.28	58.95
σ_s [dB]	7.82	6.77	8.51
$\sigma_t \sim \text{lognormal}(\mu, \sigma)$	(4.39, 0.99)	(4.00, 1.31)	(4.08, 1.13)
p-value of A-D test at $\alpha = 0.01$	0.31	0.26	0.13

3.5 Discussion

The previous power-law path-loss models in wireless mesh networks have been developed for specific locations over the short period of time thus long-term temporal variability of the communication links has not been considered. A path-loss dataset from BC Hydro MSGN for three geographical regions over the course of three months was used. Compared to the data collected with lab grade instruments, developing a path-loss model with network generated data suffers from some specific limitations. Due to the large number of communication links and number of times that the data has been captured for each communication link, filtering erroneous data can compensate the limitations of dataset. A statistical assessment was performed over path-loss dataset to identify and remove erroneous data. To characterize channel characteristic in time, Anderson-Darling goodness of fit tests were applied to the RSSI values of each individual links to make sure the RSSI values of links follow lognormal distribution in linear scale and normal distribution in logarithmic scale.

Long-term temporal and spatial variation of communication links were demonstrated in the slope and shadow fading of conventional power-law path-loss model. It was shown that previously used path-loss model cannot capture the long-term temporal variation of communication links. Additionally, path-loss parameters vary from one region to another as vegetation density and terrain morphology changes. Spatial variation in a path-loss model are due to the change in the environments of all communication links whereas the long-term temporal variation is unique for each communication link. A new path-loss model was proposed which incorporates both the long-term temporal and spatial variation of communication links. Fading in communication links was modeled by two terms to differentiate spatial and long-term temporal variation of communication links. The spatial term has a normal distribution captured from residuals of path-loss model. Therefore, in the logarithmic scale, the long-term temporal term is a normal random variable with zero mean and a standard deviation which follows a lognormal distribution. The lognormality of links' long-term temporal standard deviation distribution was verified by Anderson-Darling goodness of fit test. This way of path-loss modeling would help utility network designers to account for the long-term temporal and spatial variability of communication links separately. Based on our path-loss model, the nodes' transmit power in a mesh network can be tuned individually for different regions and time interval. As a next step of this work, capturing the long-term temporal variation of communication links in a region for the period of two years can show the trend of long-term temporal variability in communication links. Additionally, to understand the underlying reasons behind long-term temporal variation of wireless links, high resolution dataset in time is necessary.

Chapter 4: Combined Relay Node Placement-Transmit Power Adjustment Algorithm in Non-Battery Based Wireless Sensor Network

4.1 Introduction

Wireless sensor network (WSN) is a low-cost, low-power, energy efficient, ad-hoc wireless mesh network. Remote monitoring, home automation, military surveillance, border monitoring, disaster monitoring, smart utility networks (SUNs), and industrial automation are just some of broad WSNs applications. WSNs are categorized as high adaptable mesh networks with the following characteristics:

- ranged from a highly dynamic network to a static one
- covering sensing fields with different sizes, from few square meters up to a size of a province or a country
- flexible in deployment and operation phases for various network purposes. As an example, for a given sensing field, coverage, connectivity, and fault tolerance of WSNs can entirely be different based on the goal of the network.

Based on the goal and application of WSNs, design and operation of WSNs can be quite challenging in each OSI layer. In the physical layer, determining the numbers and locations of sensor and relay nodes is essential to make a fully connected network [51], [52]. Choosing a proper network power adjustment algorithm to run the network in certain power level to avoid interference as well as fulfill the desired network performance metrics (e.g., latency, throughput, network lifetime) for specific goals, are some of the design challenges [53]. To ease these challenges, relay node placement (RNP) and transmit power adjustment (TPA) are two main current solutions in designing and operating a WSN.

Typical WSNs are equipped with battery-based sensors and the amount of energy is constraint to the capacity of the power resource in sensor nodes. Energy consumption of packet transmission and reception are considerably high in sensor nodes. Therefore, relay nodes can carry the burden of collecting data from sensor nodes and prolonging network lifetime [51]. Relay nodes increase network connectivity and therefore redundancy by adding more routing paths among sensor nodes which balance network traffic [52]. If a part of WSN fails to transmit packets, data will route to base station with another path. Therefore, RNP can introduce more fault tolerant network [54]. Solutions have been proposed in the literature to find the optimum location of relay nodes to prolong network lifetime as the main goal [55], [56] and create k -connectivity ($k > 3$) for one-tiered and two-tiered WSN for homogenous and heterogenous topologies as the secondary goal. The majority of the previous works model the network as a graph and solve minimum Steiner tree problem heuristically [57]–[62]. Also, as a secondary goal, failure identification within sensor/relay nodes and federating a segmented network with minimum number of relay nodes has been the focus of some researches [63], [64]. Average node connectivity and number of deployed relay nodes are the metrics to assess different RNP algorithms [51], [52].

The RNP solution is tied with the TPA one. Applying TPA algorithms to adjust the transmitted power of each node has great impact on reducing the average consumed power of a WSN, thus prolonging the network lifetime [53]. Additionally, wireless interference between neighboring nodes affects the link quality which can be overcome by applying a TPA algorithm. Nodes that are located at the vicinity of many other nodes are more susceptible to interference [65], [66]. TPA algorithms are usually distributed algorithms which have two phases [66]. In the first phase, each node makes a node ID list, out of the nodes in its vicinity by transmitting the beacon packet and waiting for acknowledgment packet from the neighboring nodes. In the second phase, a threshold

is set by network operator based on either RSSI or link quality indicator (LQI) metrics. Finally, the transmitted power will be adjusted to meet the threshold value [67],[68],[69]. Some of the TPA algorithms repeat these two phases for multiple power level over time. Therefore, their assigned power level to each node varies [70], [71], [72]. Average transmitted power of network is one of the main metrics to assess the performance of the TPA algorithm in the literature [53].

RNP and TPA algorithms available in literature are designed for battery-based WSNs. Though, there are some WSNs e.g., SUNs are not a battery-operated WSN which are completely left out in the literature. SUNs are usually at the scale of a province or a country with IEEE 802.15.4g, as the standard for the physical and MAC layer [17], [73].

Topology is the key part in defining the optimum connectivity for any WSN. Generally, topologies in a WSN can be categorized as either homogeneous or heterogeneous nodes with a range of low to high node density [74]–[76]. Coexistence of topologies with different densities necessitates the simultaneous operation of RNP and TPA algorithms. To the best of our knowledge, no previous work assesses the simultaneous performance of RNP and TPA algorithms for various WSN topologies.

In this work, a comprehensive combined RNP-TPA algorithm is proposed for the physical layer. This algorithm is compatible with various WSN topologies. In order to reduce the susceptibility of each communication channel to interference and any possible faults, our proposed RNP-TPA algorithm is a distributed algorithm which functions based on the number of nearest neighbors (NNN) of each node. RNP part of the algorithm is responsible to increase the network connectivity by adding relay nodes within the sensing field and TPA is responsible to decrease the number of connectivity of the nodes by decreasing the transmission power of each node. A predetermined interval of NNN is defined as a reference in our algorithm. The upper bound of this

interval corresponds to the maximum number of neighbors that a node can have without experiencing interference. TPA algorithm operates based on comparing the NNN value of each node to the upper bound of the interval. The lower bound of this interval corresponds to the minimum number of neighbors that a node can have to be fault tolerant against any possible faults in the network. RNP algorithm operates based on comparing the NNN value to the lower bound of interval. Our combined algorithm is applied to WSNs with different node density and network topologies. To summary, the objectives of this chapter are:

- To propose a combined RNP-TPA algorithm in physical layer which operates based on NNN metric for each node. Network lifetime is not considered in this algorithm; thus, it can only be applied to non-battery based WSNs with any network topology and node density.
- To demonstrate domain of applicability of the combined algorithm based on node density and quantified topology irregularity
- To evaluate the algorithm performance based on two main metrics in WSN i.e., the average transmitted power of nodes and number of added relay nodes.
- To assess the effectiveness of our proposed algorithm for each topology by comparing average and standard deviation of nodes' NNN before and after applying the algorithm. This can clarify if NNN variations is fitted within the predetermined NNN range. The rationale behind it is to measure the extent that combined algorithm can fit the width of NNN distribution to the predetermined range of NNN for each topology irregularity.

The remainder of this chapter is organized as follows. Concept of NNN and the proposed RNP-TPA algorithm are introduced in Section 4.2. Methodology of the simulation and the results are covered in Sections 4.3 and 4.4, respectively. Finally, Section 4.5 concludes the chapter.

4.2 Concepts

This section provides the concept of NNN as a key metric to implement the combined RNP-TPA algorithm. The operation of the combined RNP-TPA algorithm based on the NNN is explained further.

4.2.1 Number of Nearest Neighbors

In an IEEE 802.15.4 standard MAC layer, every node has a table node ID which contains the information of its neighboring nodes [77]. Every row of this table corresponds to the nodes that can have a two-way communication with the node under study. NNN is the number of rows in the table node ID. Given the topology of WSN, NNN can get a range of low to high values. NNN is used in TPA and RNP algorithms separately and is used here as a decision-making metric for our combined algorithm.

Given uniform transmission power for all nodes, in an irregular network topology, nodes are located far from each other where only one or two edge-disjoint connectivity from nodes to base station exists. Therefore, network is not fault-tolerant, and any node/communication link fault would lead to a segmented network. To find the location of required relay nodes in the network topology, L_{NNN} is defined as the lower bound of NNN. Nodes with NNN less than L_{NNN} are vulnerable to faults and their connectivity needs to be increased. In a regular network topology, nodes are located uniformly with the same number of neighbored nodes. Any alteration in the transmission power will lead to rapid changes in the NNN value of each node. U_{NNN} is defined as the upper bound of NNN. Nodes with NNN value greater than U_{NNN} are more susceptible to

interference [78], [79]. We call $L_{NNN} - U_{NNN}$ the predetermined NNN interval set as an input in the RNP-TPA algorithm. Based on the goal of the network, the number of tiers, level of expected connectivity, the coverage of sensing field, and etc. $L_{NNN} - U_{NNN}$ needs to be defined by the network operator. These bounds can get different values in various parts of the network based on the topology.

Here, NNN is a figure of merit which can be incorporated in both RNP and TPA algorithms. In the low node density areas within WSNs, RNP algorithms metric is the number of disjoint connectivity between individual nodes and base stations. In low node density areas of a WSN, NNN metric can represent the number of connectivity. The only exception is a WSN topology with two clusters of nodes that are connected through one communication link (bottle neck) [51], [60]. In the high node density areas, TPA algorithm metric is a link quality indicator such as ETX. NNN can represent the ETX metric in high node density areas[67].

4.2.2 RNP-TPA Algorithm

Figure 4.1 shows a block diagram of our proposed combined RNP-TPA algorithm. The combined RNP-TPA algorithm proposed in this paper is based on the predetermined interval of NNN in each node. Comparing the value of NNN for each node with the U_{NNN} and L_{NNN} , our algorithm decides whether to use RNP or TPA or both for each node. Nodes with NNN less than L_{NNN} are prone to be segmented from the rest of network due to the possible faults within the nodes and communication links. Therefore, the RNP part of the combined algorithm adds relay nodes to increase nodes' connectivity up to the L_{NNN} . Nodes with NNN greater than U_{NNN} are susceptible to interference. In this case, the TPA part of the combined algorithm decreases the node transmission power level so that its NNN is less than U_{NNN} . It is assumed that all the nodes have a uniform transmission power and transmit at their maximum power level. Nodes

connectivity can be increased only by RNP algorithm, therefore, RNP should be run prior to TPA in the combined algorithm. Placing TPA prior to RNP would increase the number of iteration in algorithm.

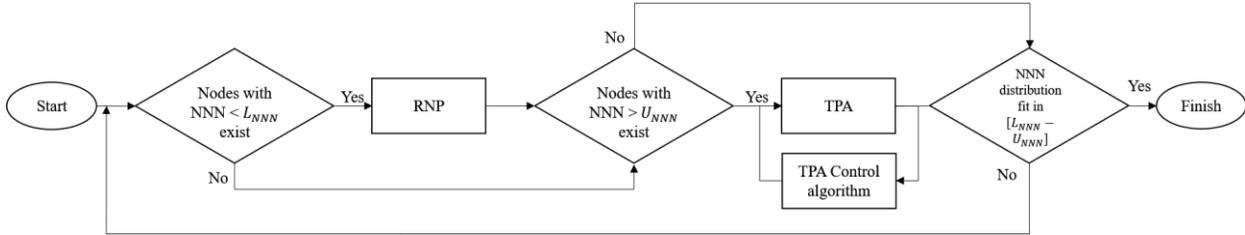


Figure 4.1. Proposed RNP-TPA algorithm

The step by step procedure of our algorithm is as follows: It starts with identifying all the nodes with NNN value less than L_{NNN} . Next, the proposed RNP algorithm runs for all the nodes that their connectivity needs to be increased. Once the RNP algorithm is finished, network connectivity, and consequently NNN values for each node are updated. At this point, combined algorithm identifies nodes with NNN values higher than U_{NNN} are identified by the algorithm. Next, the TPA algorithm runs for all the nodes so their power can be reduced to lower the probability of interference. Finally, nodes NNN values get updated by the algorithm again and if nodes with NNN value out of $L_{NNN} - U_{NNN}$ are found, the algorithm converges when node NNN distribution meets the following requirements: 1) Average of NNN distribution lies within the predetermined range of NNN and 2) standard deviation of NNN distribution gets a smaller value than the length of predetermined NNN. The convergence condition implies that at least 68% of the node NNN values after applying the algorithm, lies within the nodes predetermined NNN range. The length of predetermined interval is the threshold set by the authors and is shown by $t_{\sigma_{NNN}}$. Based on the goal and usage of WSN, network operator might set different value for this threshold. Note that if no nodes can be found with NNN values below (beyond) the L_{NNN} (U_{NNN}) the algorithm will skip the RNP (TPA) part.

In summary our proposed RNP-TPA algorithm are applied to the nodes with NNN values out of the predetermined range. The goal is to change the NNN value to be within the predetermined interval. The complexity of combined algorithm equals to the multiplication of RNP and TPA complexities. RNP and TPA block diagrams and their complexities are explained in the next subsection.

4.2.3 RNP Algorithm

The RNP algorithm is designed to increase the connectivity of nodes with NNN less than L_{NNN} . RNP algorithm which is a greedy heuristic algorithm implemented based on the convex hull algorithm mentioned in [63]. In our work, the convex hull algorithm is adjusted within the following three steps to operate with NNN metric:

- Nodes with NNN less than L_{NNN} are called candidate nodes. RNP algorithm starts with selecting a candidate node with the minimum number of NNN and attempts to create a convex hull around the candidate node. The convex hull forms a polygon shape tier around the candidate node. Determining the size of convex hull and correspondingly number of unengaged nodes (including the candidate node) inside the first convex hull is of great importance. If the number of unengaged nodes is less than three the connectivity of the candidate node cannot be increased for some specific topologies. The RNP algorithm starts creating a convex hull with all the nodes that are located at the vicinity of candidate node within its communication range. If the number of unengaged nodes inside the created convex hull is smaller than three, the RNP algorithm would increase the search radius, to an extent that three unengaged nodes remains in the first-tier convex hull.
- Every three neighbors in a convex hull which are not located in one line can form a triangle [80]. Nodes which creates the convex hull might also be part of the candidate nodes. In this

scenario, for every triangle, if NNN value of at least one of three nodes is less than L_{NNN} , RNP will find the Fermat point of the triangle to add a relay node. To determine the location of the Fermat point, RNP creates the smallest rectangle which includes the triangle vertices. A regular mesh grid is created within the rectangle area. The diagonal distance of two adjacent grid vertices in this mesh grid is set to quarter of the nodes communication range which can be calculated from path-loss model. Once the Fermat point is calculated, if the distance between each vertex of the triangle and Fermat point is larger than the transmission range of nodes, RNP adds intermediate points so that every node in triangle can communicate with its Fermat point. Intermediate points are added in a straight line which connects Fermat points to the vertices of triangle. The number of intermediate points is determined by dividing the distance between the Fermat point and each vertex over the node communication range. Once all the Fermat points and intermediate points in the first convex hull are identified and replaced with relay nodes, RNP updates network parameters such as NNN, and node distances.

- In the final step, all the unengaged nodes in the convex hull in addition to the added relay nodes in the previous step, form another convex hull which is smaller than the previous one. The RNP repeats all the previous steps and updates the network parameters again. For every candidate node in each convex hull, NNN is at least L_{NNN} and it's fully connected to the inner tier convex hull. This phase of RNP algorithm continues until only three unengaged nodes remain in the n^{th} convex hull. In our proposed RNP, relay nodes are added from the first to n^{th} tier of convex hull toward the core of the convex hull where the candidate node is located.

The RNP algorithm continues to an extent that all candidate nodes connectivity increased by placing relay node around them. Although, the goal of RNP algorithm is to increase the nodes connectivity to L_{NNN} , in some network topologies this cannot be achieved in one iteration. The

convex hull algorithm is calculated by Graham scan algorithm with the complexity of $O(n \log n)$. n is the number of nodes which convex hull are created with [81]. Our proposed RNP algorithm have the complexity of $O(N_t N_c (\log N_c + c))$ [63]. N_t is the number of tier that convex hull are created for each candidate node, N_c is the number of nodes that comprises each convex hull and c is the number of vertices in the mesh grid that RNP searches for Fermat point.

4.2.4 TPA Algorithm

The TPA algorithm is responsible for reducing the power of nodes with NNN greater than U_{NNN} . The input of the TPA algorithm is the table node ID of all the nodes in WSN including RSSI values of neighboring nodes. The TPA algorithm starts with calculating NNN values and range of coverage for each node based on the free space path-loss model. TPA algorithm applies the following steps to every node with NNN greater than U_{NNN} (candidate nodes):

- TPA starts with candidate nodes with the highest NNN values. it calculates distance of the candidate node to all of its neighbors with their NNN values and sort them based on distance in descending order in a list. Distance of each node can be calculated based on the RSSI value and path-loss model.
- TPA strategy is to decrease power of the candidate node to have U_{NNN} neighbors, though based on the NNN of neighbors, this may or may not happen:
 1. If all neighbors have NNN greater than U_{NNN} , TPA starts selecting the neighbors from the bottom of the list and uses the distance of U_{NNN}^{th} row to calculate the power threshold of candidate node. Therefore, candidate nodes' power cannot cover any distance bigger than the distance of U_{NNN}^{th} row. In this case, number of neighbors is equal to U_{NNN} .

2. At least a neighbor with NNN less than U_{NNN} exists in the list. After finding the neighbor with the minimum NNN, starting from the bottom of the list, TPA compares the location of these rows with the U_{NNN}^{th} row. The row with higher location in the list from the bottom is selected to calculate the threshold power of the candidate node. In this case, number of selected neighbors of the candidate node may be more than U_{NNN} .

The TPA algorithm continues to an extent that all candidate nodes connectivity decreased by decreasing nodes transmission power. Although, the goal of TPA algorithm is to decrease the nodes connectivity to U_{NNN} , in some network topologies this cannot be achieved in one iteration. While decreasing the power, if NNN value of any neighbor nodes goes below U_{NNN} , the TPA control algorithm stops decreasing the power of nodes to avoid possibility of islanding neighboring nodes. In this situation, NNN value of nodes are greater than U_{NNN} , but the need for adding relay nodes is eliminated. Therefore, the combined RNP-TPA algorithm runs for the fewer number of iterations. The complexity of TPA algorithm is $O(N_c N_{C_{NNN}})$. N_c is the number of candidate nodes and $N_{C_{NNN}}$ is the number of nearest neighbours for each candidate node.

4.3 Methodology

The combined RNP-TPA algorithm is applied to several synthesized WSNs with a range of topology irregularities and node densities simulated with MATLAB. Simulation parameters are depicted in Table 4.1 [82]. In the next three subsections, the detailed topology irregularity, RNP and TPA algorithms are explained.

4.3.1 Topology Irregularity

An irregularity index λ is introduced to measure the irregularity of the synthesized network topology. WSN topology can vary from a regular topology ($\lambda = 0$) in which nodes are located in a mesh grid, to a completely irregular network ($\lambda = 1$) in which nodes form clusters with various

densities. To synthesize a WSN with random topology, first number of nodes N in the sensing field is selected. Then, a uniform mesh grid comprises of N nodes with x and y coordinate is created. This guarantees that all sensor nodes are located at the vertex of similar rectangles. To create an irregular network, an error is defined and added to the coordinates of all nodes in the uniform mesh grid. The diagonal of rectangle D which is the distance between two adjacent nodes, is selected as the maximum standard deviation of error that can be applied to the nodes.

Then, a normally distributed random error with the mean of zero and standard deviation of $\lambda * D$ is applied to both x and y coordinates. Thus, $\lambda = 0$ implies no error i.e., the uniform grid and $\lambda = 1$ implies the most irregular network in our simulation. After applying error, the network nodes' locations get updated. Next, power of each node P_0 is selected. Finally, given the free space path-loss model, sensitive of nodes and shadow fading affect the coverage range for each node is calculated. To remove the effect of random error which is applied to each topology, the simulation runs over 20 times for each topology and all the desired metrics are averaged values.

No information has been reported in previous research regarding the network irregularity and its implementation. The proposed intuitive simulation method can be a motivation for further investigation and research to improve the implementation and quantification of the network irregularity.

Table 4.1. Simulation parameters

Area [Km^2]	$5 \times 5 = 25$
Diagonal distance between two vertices [m]	{772, 467, 368, 317, 280, 259, 241, 225, 212, 199}
Number of Nodes	{125, 250, 375, 500...1250}
Density [N/Km^2]	{5, 10, 15, 20..., 50}
Irregularity Index	[0 -1]
Nodes Power [dBm]	{25, 30}
Path-loss Model	Free Space
Sensitivity [dBm]	-95
Shadow fading [dB]	6.85
L_{NNN}	3
U_{NNN}	5
$t_{\sigma_{NNN}}$	2

4.4 Results

In this section the simulation results are presented in three subsections. In the first subsection the domain of applicability of the combined RNP-TPA algorithm is presented for various topology irregularities and node densities. In the second subsection, the evaluation of our algorithm is discussed based on nodes' average power and number of added relay nodes. Last subsection shows the effectiveness of the algorithm in adjusting the NNN distribution within the predetermined interval by comparison of the average and standard deviation of NNN before and after applying the algorithm.

4.4.1 Domain of Applicability

Given P_0 and for a range of node densities and topology irregularities, Table 4.2 and 4.3 indicate whether RNP, TPA or combined RNP-TPA algorithm is used. It is observed that:

- In Table 4.2, for the regular, grid shape WSN topology ($\lambda=0$) and low node densities where neighbors are not at the coverage range of each other, only RNP is applied. As node density increases, more nodes are located at the coverage range of each other, thus both RNP and TPA are required. For higher node densities only TPA are required. In table 4.3, where the coverage range of the nodes are longer, only TPA is required to tune WSN.
- As the network becomes more irregular in both tables 4.2 and 4.3, for the majority of node densities, both RNP and TPA algorithms are required to tune the network. In irregular topologies, network nodes form clusters with nonuniform densities which necessitate adding relay nodes for some and reducing nodes power for the other ones.
- In Table 4.2 with lower node coverage, the algorithm starts with applying RNP for low node densities. As node density increases both RNP and TPA algorithms are required. For longer node coverage in table 4.3, the algorithm applies both RNP and TPA initially, and by increasing node density only TPA is applied. Two extreme cases within the domain of applicability of the algorithm are distinguished: 1) Low node power and minimum node density, and 2) high node power and maximum node density. Independent of the irregularity, RNP and TPA are required respectively for the former and latter cases to tune the network.

Table 4.2. Applicability of RNP-TPA, P₀ = 25 dBm

Density [N/Km^2]	5	10	15	20	25	30	35	40	45	50
λ										
0	RNP	Both	Both	Both	TPA	TPA	TPA	TPA	TPA	TPA
0.5	RNP	RNP	Both							
1	RNP	RNP	Both							

Table 4.3. Applicability of RNP-TPA, P₀ = 30 dBm

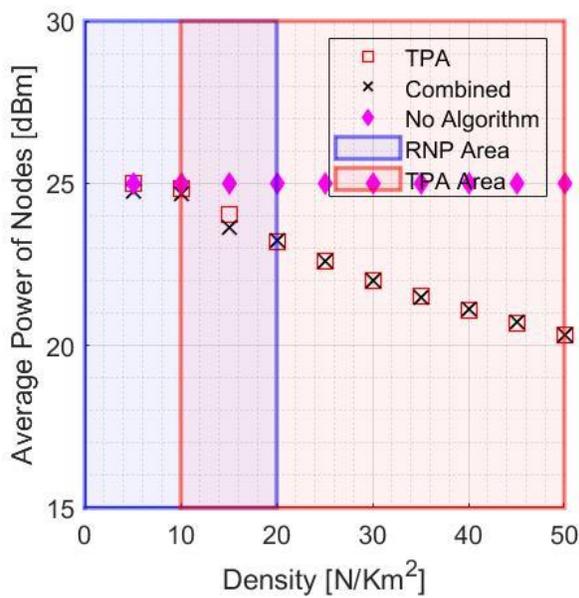
Density [N/Km^2]	5	10	15	20	25	30	35	40	45	50
λ										
0	TPA	TPA								
0.5	Both	Both	Both	Both	TPA	TPA	TPA	TPA	TPA	TPA
1	Both	TPA								

4.4.2 Evaluate Algorithm Performance

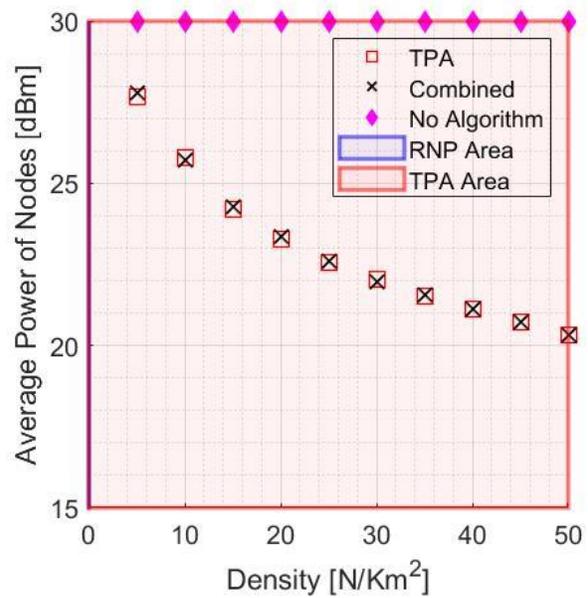
To evaluate the performance of the algorithm, variations of the average power of nodes and number of added relay nodes which are two important metrics in the physical layer, are shown for various topology irregularities and node densities. Based on the result in the previous subsection, domain of applicability of TPA and RNP algorithms are depicted with red and blue area, respectively. In figures 4.2 and 4.3. the border of these two areas depends on λ and P_0 . Domain of applicability of the combined RNP-TPA algorithm is the overlapped part of these two areas.

Figure 4.2 shows the average nodes' power versus node density for various λ and P_0 when applying TPA and combined RNP-TPA algorithms. In all cases, as node density increases, average power of nodes decreases. In the overlapped domain, combined algorithm has a better performance in reducing power of nodes compared to the TPA since RNP increases the connectivity of the nodes

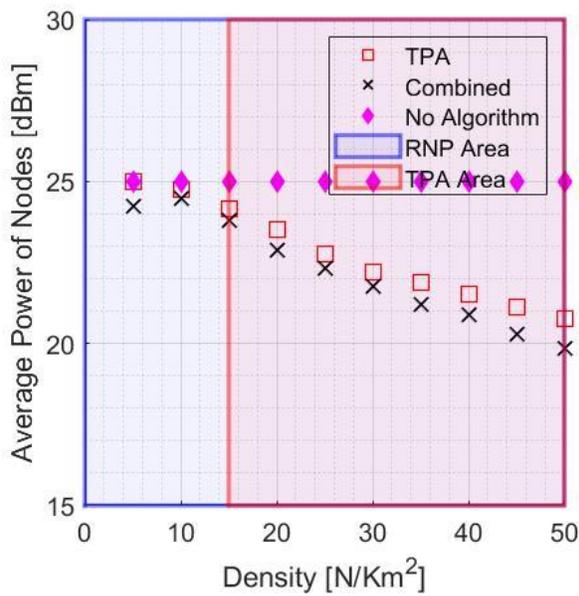
prior to applying the TPA. In the TPA domain, the performance of combined and TPA algorithm are similar since no addition relay nodes is required. In each case, the amount of decrease in average node power are independent of node initial power and completely depends on node density and network irregularity. For higher value of P_0 , the amount of decrease in average power of nodes are greater. For a given node density, as network becomes more irregular, the amount of decrease in power gets a smaller value.



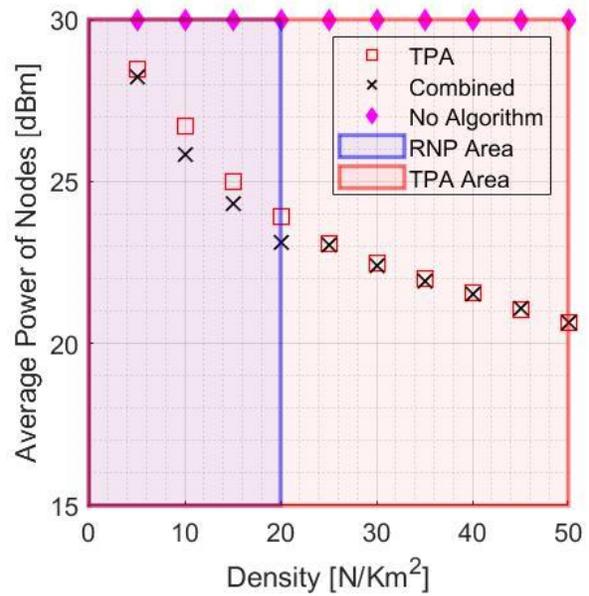
$\lambda=0$ & $P_0=25$ dBm



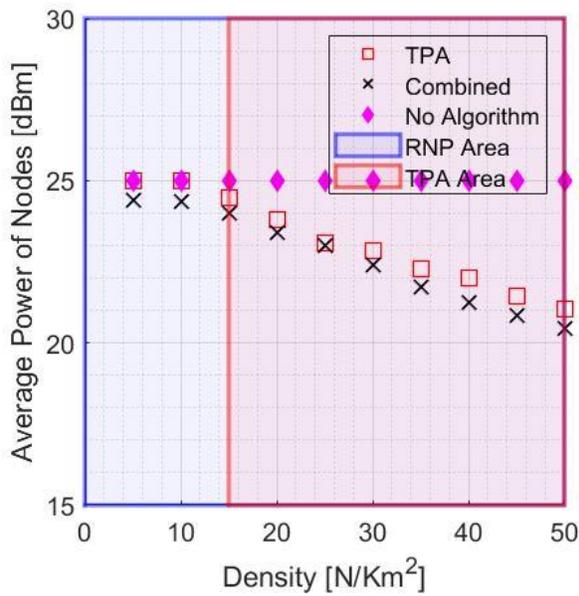
$\lambda=0$ & $P_0=30$ dBm



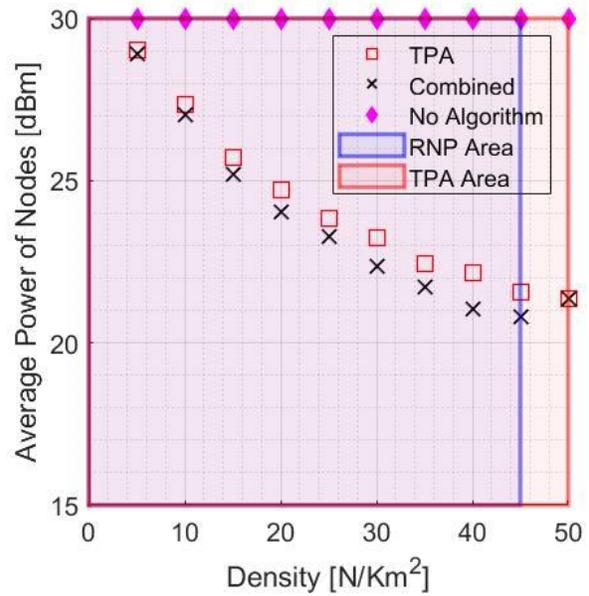
$\lambda=0.5$ & $P_0=25$ dBm



$\lambda=0.5$ & $P_0=30$ dBm



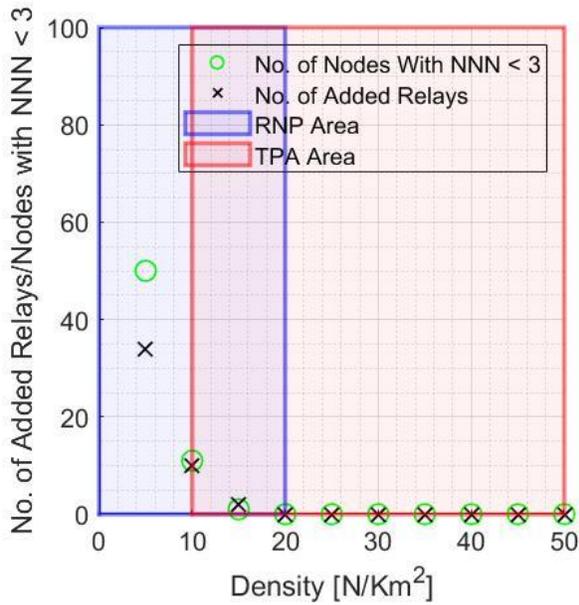
$\lambda=1$ & $P_0=25$ dBm



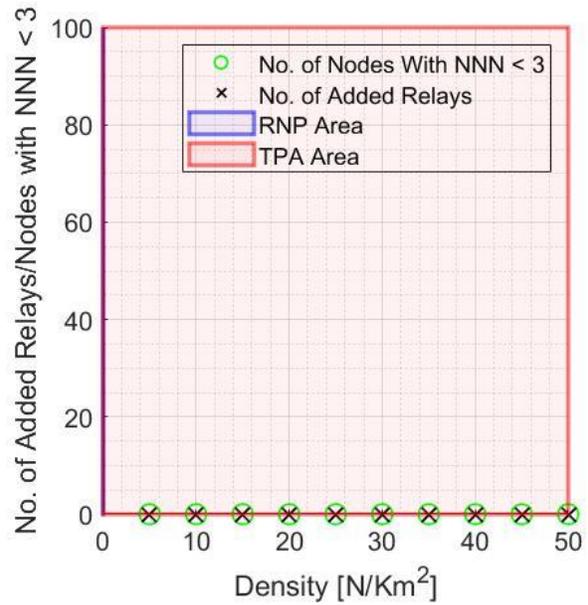
$\lambda=1$ & $P_0=30$ dBm

Figure 4.2. Average power of nodes vs. node density for various λ and P_0

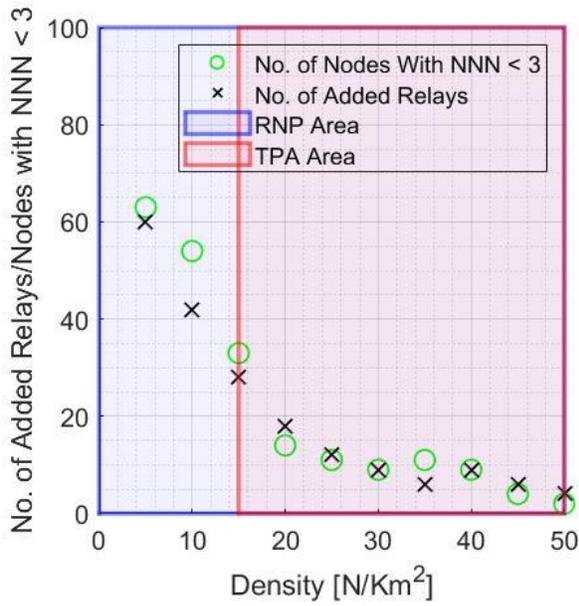
Figure 4.3 shows the number of added relay nodes versus node density for various λ and P_0 . For each case, the number of nodes with NNN less than L_{NNN} are shown. In all cases, as node density increases the number of added relay nodes decreases. After RNP domain, number of added relay nodes is zero. For a given node density as topology irregularity increases, the number of added relay nodes increases.



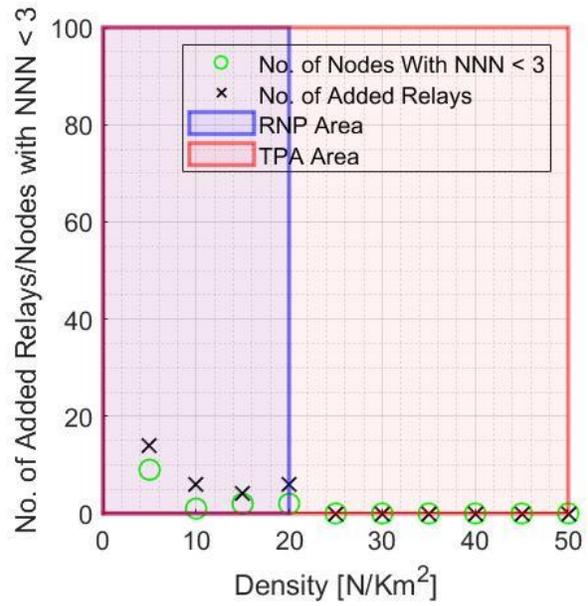
$\lambda=0$ & $P_0=25$ dBm



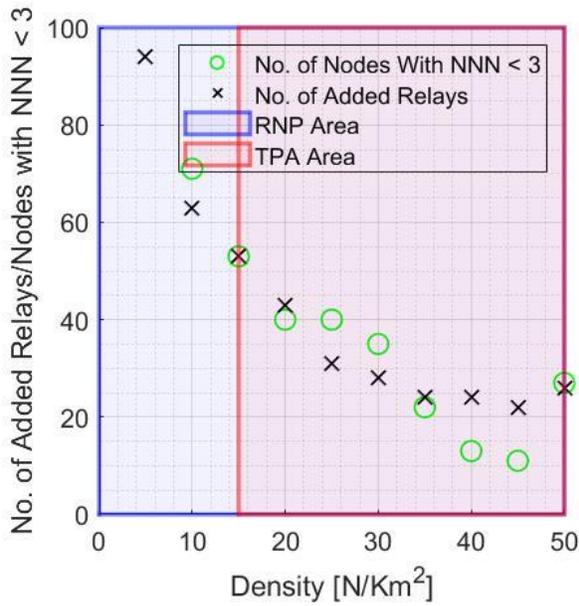
$\lambda=0$ & $P_0=30$ dBm



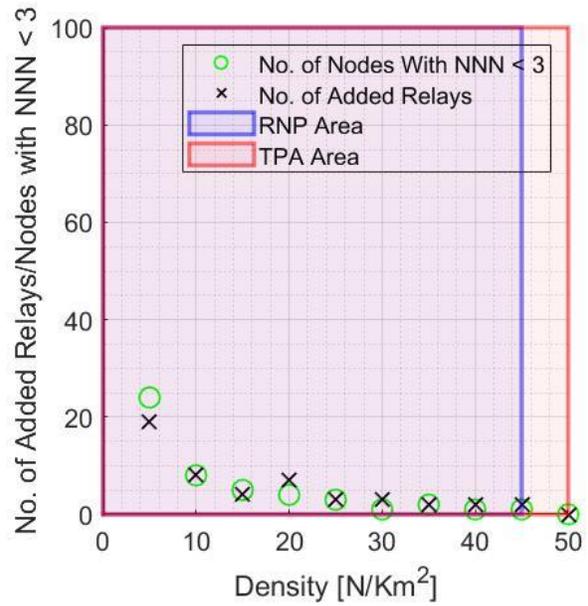
$\lambda=0.5$ & $P_0=25$ dBm



$\lambda=0.5$ & $P_0=30$ dBm



$\lambda=1$ & $P_0=25$ dBm



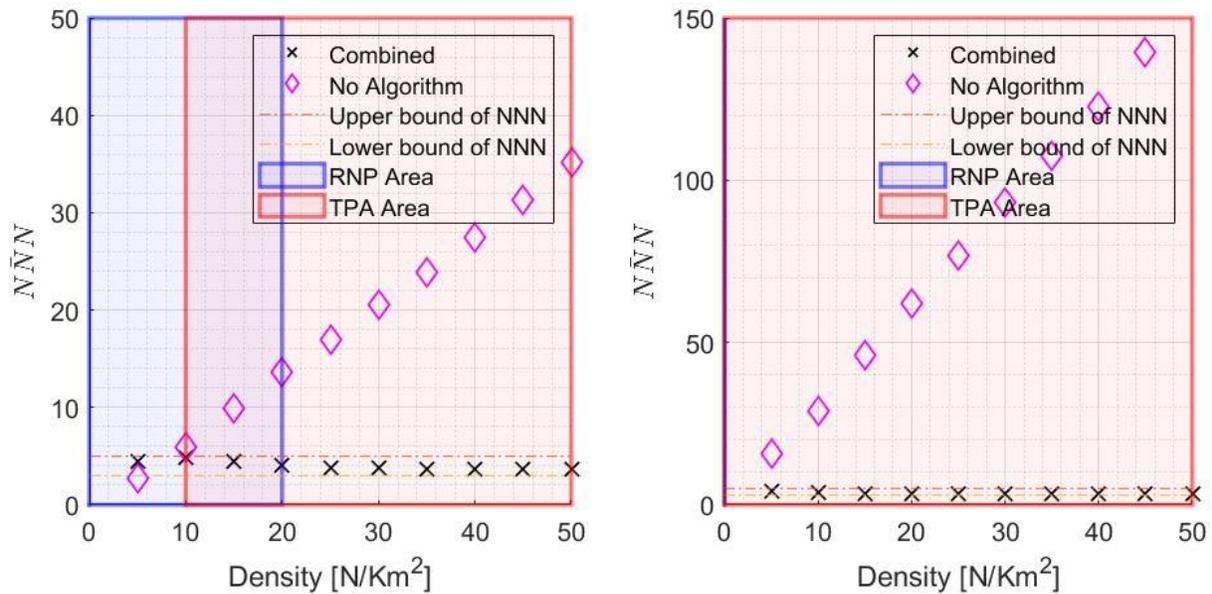
$\lambda=1$ & $P_0=30$ dBm

Figure 4.3. Number of added relay nodes vs. node density for various λ and P_0

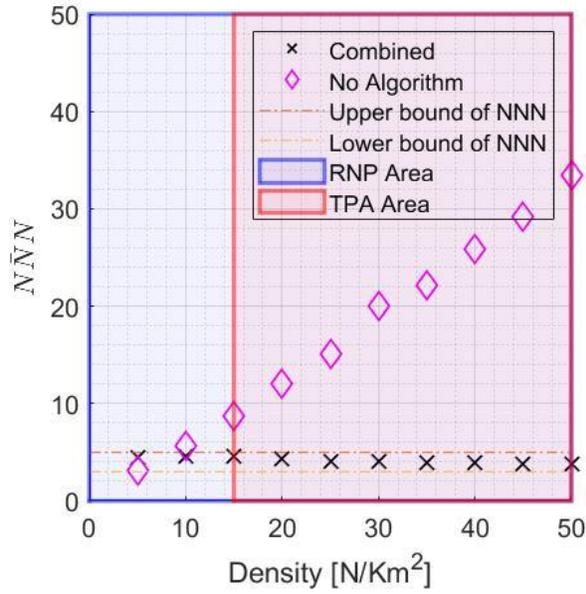
4.4.3 Effectiveness of the Algorithm

The strategy of the proposed algorithm is to alter nodes' NNN values so that NNN distribution fits in the predetermined range of NNN 3-5 after applying the combined RNP-TPA. This is illustrated by comparing the average and standard deviation of NNN values (\overline{NNN} and σ_{NNN}) with predetermined range of NNN before and after applying the combined RNP-TPA algorithm in figures 4.4 and 4.5.

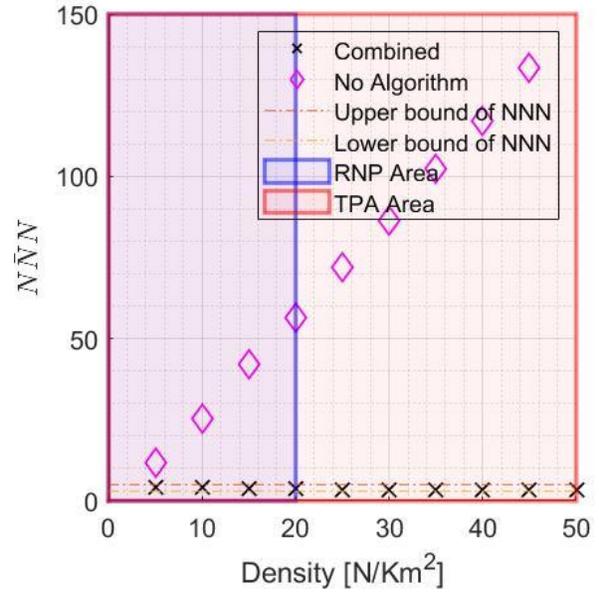
In figure 4.4, before applying the combined algorithm, independent of λ and P_0 , as node density increases, \overline{NNN} increases. In the network with higher P_0 the coverage range of each node is more, thus \overline{NNN} is higher. For a given node density as λ increases, \overline{NNN} is reduced. After applying the algorithm, \overline{NNN} lays within 3-5. Figure 4.5 shows the σ_{NNN} before and after applying the combined algorithm for various λ and P_0 . Before applying the algorithm, as node density increases, σ_{NNN} increments, though after applying the algorithm, σ_{NNN} reduces to a value below the length of predetermined range ($t_{\sigma_{NNN}}$) and remains constant independent of network irregularity.



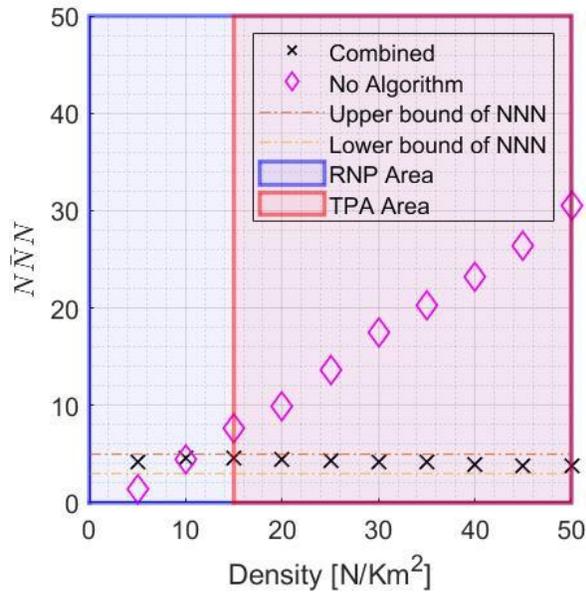
$\lambda=0$ & $P_0=25$ dBm



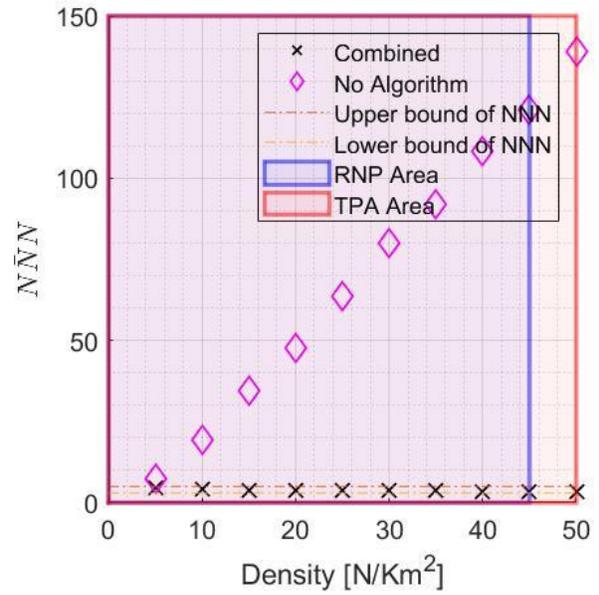
$\lambda=0$ & $P_0=30$ dBm



$\lambda=0.5$ & $P_0=25$ dBm



$\lambda=0.5$ & $P_0=30$ dBm

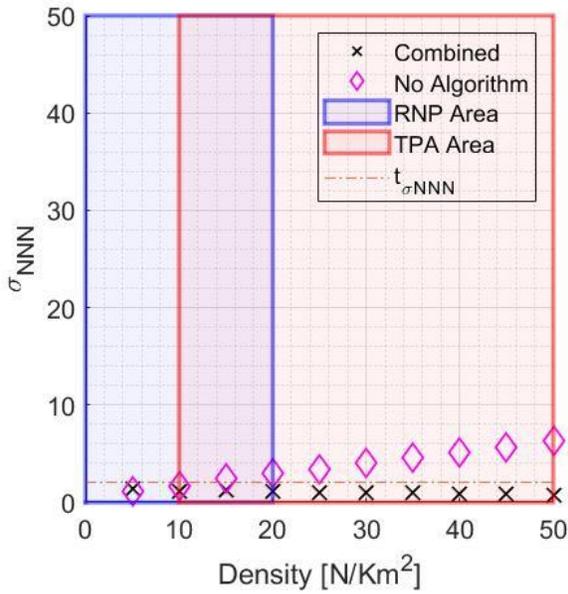


$\lambda=1$ & $P_0=25$ dBm

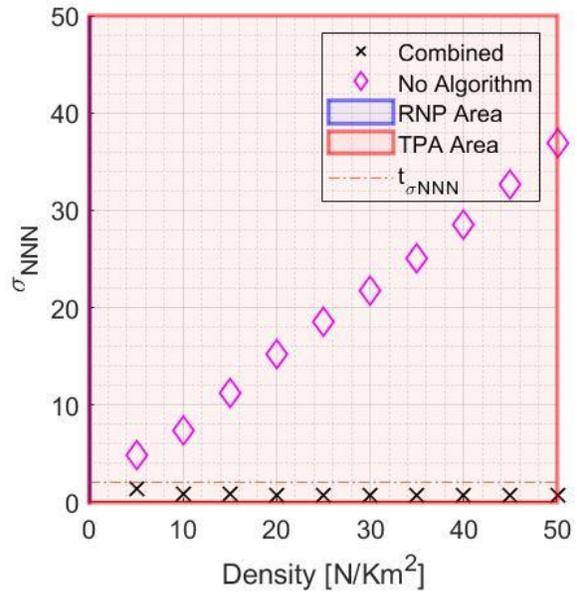
$\lambda=1$ & $P_0=30$ dBm

Figure 4.4. \overline{NNN} vs. node density for various λ and P_0

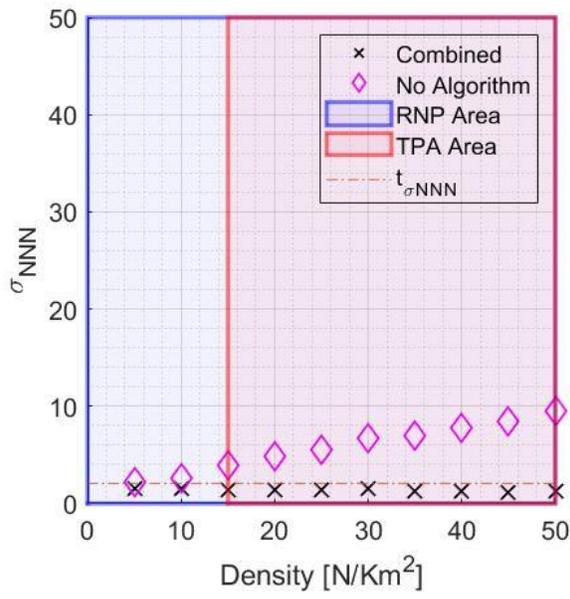
The combined algorithm runs for as many iterations as needed to change the distribution of NNN in a way that : 1) \overline{NNN} gets a value within the 3-5 which is the predetermined range of NNN and 2) σ_{NNN} gets a value smaller than 2 which is the length of predetermined range. The combined algorithm meets the requirement in the first iteration. The combined RNP-TPA algorithm convergence for each node density and network topology in this simulation is equal to one. Decreasing the requirement for nodes standard deviation would increase the number of cycles that RNP-TPA needs to be executed.



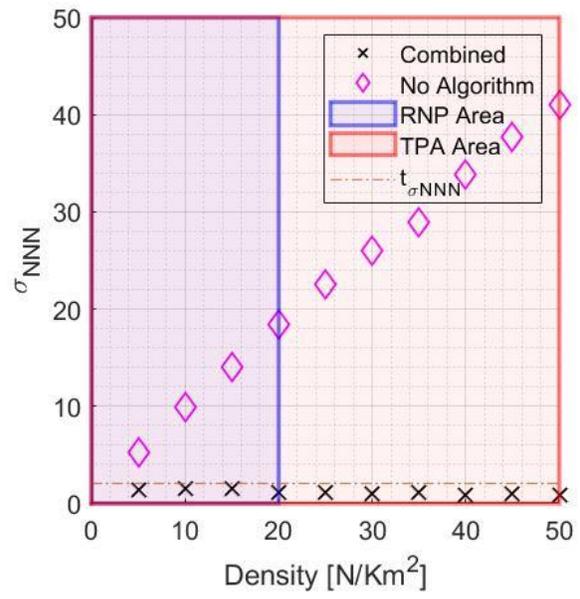
$\lambda=0$ & $P_0=25$ dBm



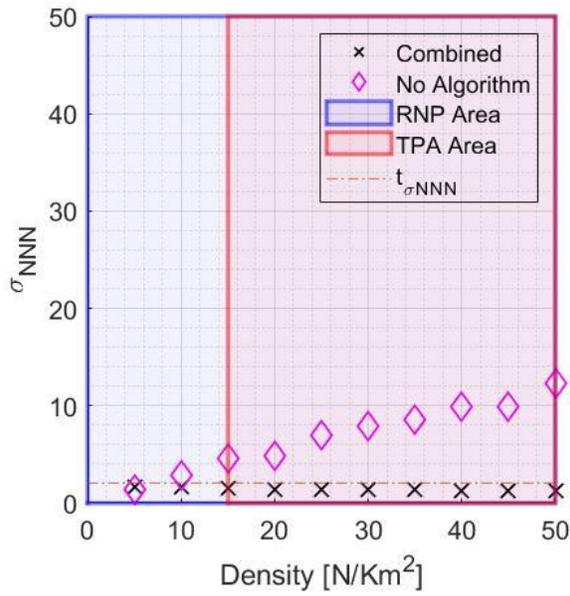
$\lambda=0$ & $P_0=30$ dBm



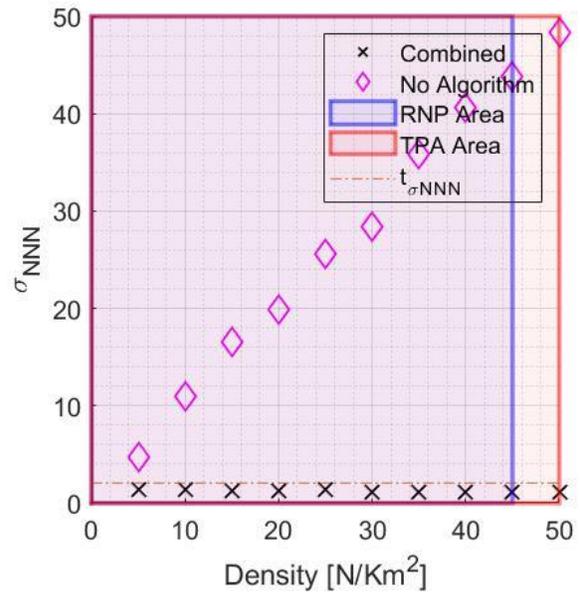
$\lambda=0.5$ & $P_0=25$ dBm



$\lambda=0.5$ & $P_0=30$ dBm



$\lambda=1$ & $P_0=25$ dBm



$\lambda=1$ & $P_0=30$ dBm

Figure 4.5. σ_{NNN} vs. node density for various λ and P_0

4.5 Discussion

Numerous works have been done to tune and optimize WSNs in physical layer by either adjusting the power of nodes to reduce the total consumed power or adding relay nodes to increase the connectivity of the nodes. To perform both tasks together, in this work we proposed a combined RNP-TPA algorithm based on the nodes' NNN value. The network lifetime was not considered in our algorithm which makes it suitable for the non-battery based WSN such as BC Hydro multi-service grid network. A predetermined range of 3-5 was defined for nodes' NNN values. Beyond this range nodes are susceptible to interference and below this range nodes are prone to lose their network connectivity in case of network failure.

Our algorithm decides to operate either RNP or TPA or both for nodes with NNN values out of the predetermined range. The algorithm was evaluated for various network topologies quantified with topology irregularity index. The domain of applicability of the combined RNP-TPA algorithm was obtained for various network topologies and node densities. Given maximum transmitted power in all nodes, regular topology requires TPA algorithm while combined RNP-TPA algorithm is required more in irregular topologies. Additionally, low node densities and high node densities WSNs require RNP and combined RNP-TPA or TPA algorithm, respectively. The algorithm was evaluated by average node power and number of added relay nodes. For a given irregularity index and node density, we showed how applying our algorithm can decrease the consumed transmitted power of nodes. We also showed the effect of topology irregularity on the number of added relay nodes required to be placed in a WSN.

The effectiveness of our algorithm was demonstrated by comparing the average and standard deviation of nodes' NNN values with the predetermined range of NNN. The algorithm runs until the average value of NNN distribution lies within the NNN predetermined range and the standard

deviation of it gets a smaller value than the length of the NNN predetermined range. Based on the goal and application of networks, the predetermined range may need to be changed, thus the combined RNP-TPA algorithm might be run for more than one iteration. In future, to assess WSN performance, routing algorithm in the network layer needs to be integrated. Implementing constrained location in adding relay nodes and determining a standard approach to measure topology irregularity are possible other steps.

Chapter 5: Conclusion and Future Work

5.1 Conclusion

Wireless links in large scale non-battery based wireless mesh networks such as smart utility networks are susceptible to interference in high density areas and are prone to get segmented from the network in low density areas. Overcoming those problems in physical layer can be done by adjusting nodes power and adding relay nodes to the sensing field which require investigation in path-loss model of wireless links. To tune such network, here we sought: 1) to statistically assess the accuracy of ordinary least square and method of moments methods in estimation of slope for short-range power-law path-loss models in the presence of distance error and shadow fading, 2) to develop a power-law path-loss model which incorporate the spatial and long-term temporal variability of communication links in sub-urban environment, and 3) to develop a distributed, combined relay node placement-transmit power adjustment algorithm based on power-law path-loss model for non-battery based wireless sensor networks.

The accuracy of estimated slope in short range power-law path-loss model is of great importance. Due to the inaccuracy in finding the precise location of nodes in WSNs path length of communication links are prone to distance error which results in error in independent variable. Additionally, shadow fading creates error in dependent variable which affects the estimated slope of path-loss model. In the short-range wireless communication with simultaneous presence of shadow fading and distance error, path-loss data is less correlated, therefore, path-loss slope cannot be estimated accurately. Here, the underlying statistics performance of two OLS and MoM methods in slope estimation of short-range power-law path-loss model in the presence of distance error and shadow fading was investigated.

First, behavior of MoM and OLS was investigated for wireless communication below the distance span of 100 metres. It was shown while OLS method fails to estimate the path-loss slope in the presence of distance error MoM give accurate slope estimation. Second, applicability of MoM for a given path-loss dataset was determined by proposing “method selector index”. Method selector index is calculated based on the error in estimated slope and varies with the slope of path-loss model. This index decides whether to use OLS or MoM based on the standard deviation of distance span and distance error for a given path-loss dataset, so the slope error is below 0.1. Third, the effect of shadow fading in estimation of path-loss slope with MoM was considered. It was shown that increase in shadow fading would result in decreasing the correlation of RSSI and path length in a given path-loss dataset which affects the precision of MoM in estimating the slope. To determine the accuracy of MoM with the presence of shadow fading, thresholds were set for the covariance of given path-loss dataset. Violation of these thresholds would result in false estimation of slope by MoM.

Existence of a comprehensive and accurate path-loss model which incorporates the long-term temporal and spatial variability of wireless links is vital to tune WMNs in physical layer. All the previous path-loss models were developed for a specific location over a short period of time. Acquiring path-loss dataset for various locations over long time with lab grade instruments are costly and time inefficient. Therefore, a network generated path-loss data from BC Hydro multi-service grid network were used for three different regions of Richmond, Kamloops and Terrace with various vegetation density and terrain morphology over the course of three month. Two main problem of BC Hydro dataset were, precision error in calculating wireless links path length and biased in network data for the links with higher RSSI values. Due to vast number of wireless links

and number of times each link was measured, all the confinement of network generated dataset can be resolved by filtering.

A statistical analysis was done to see the scale of BC Hydro MSGN dataset, characterize fading of wireless channels and distribution of wireless links in time. It was seen that amplitude of individual wireless links over time have lognormal distribution in linear scale and normal distribution in logarithmic scale. Anderson Darling goodness of fit test was applied to individual links to verify the lognormality of links distribution in linear scale. It was observed that 70 percent of links amplitude follow lognormal distribution in time. It was also shown that shadow fading is the source of fading within BC Hydro MSGN, based on signal bandwidth, period, channel bandwidth and receiving time of receiving nodes. To indicate the existence of long-term temporal and spatial variation of wireless links, slope and shadow fading of all communication links in each region, calculated daily over the 3-month period. To figure out the most accurate method for estimating path-loss parameters, method selector index has been applied to the dataset. Based on the result of this test ordinary least square method is used to estimate path-loss parameters.

A new path-loss model which incorporate long-term temporal and spatial variation of wireless links was proposed. It was demonstrated that spatial variation of wireless links is due to change in environments of all links in a region and have normal distribution with zero mean and a standard deviation which calculates from the residuals of path-loss model. The long-term temporal variation of individual links is a normal distribution with mean of zero and a standard deviation which follows a lognormal distribution. distinguishing the long-term temporal and spatial fading in power-law path-loss model would help network operators to prevent any possible segmentation in the network. Path-loss model is an essential metric to tune WMNs in physical layer.

Tuning a WSN in physical layer includes, adjusting nodes power in high density clusters to remove interference among them and placing relay nodes in low density clusters to increase connectivity and make the network fault tolerant. Here a combined relay node placement-transmit power adjustment was proposed based on the number of nearest neighbors of each node in WSNs. As network lifetime of the nodes were not considered, the algorithm is only suitable to non-battery based WSNs. Given a WSN with nodes transmitted power at their maximum level, a predetermined range of NNN was defined as a metric to joint two parts of the algorithms. Any nodes with NNN value below the predetermined range needs relay node to increase their connectivity and any nodes with NNN value above the range needs a reduction in their transmitted power. The goal of combined RNP-TPA algorithm is to fit nodes NNN distribution to predetermined range of NNN. The range of [3-5] was set here for the predetermined range of NNN.

To show the domain of applicability of algorithm for various topology and node density, irregularity of network topology compare to the mesh grid was quantified. It was demonstrated that TPA would be required algorithm to tune a regular mesh grid and as the network become more irregular the combined RNP-TPA algorithm is more applicable. Additionally, for a WSN with given node density and irregularity, the combined algorithm was evaluated by the amount of reduction in nodes transmitted power and number of added relay nodes. Finally, the effectiveness of our algorithm was demonstrated, by comparing average and standard deviation of nodes NNN distribution, with the predetermined range. The algorithm stops working when average nodes NNN distribution lies within the predetermined range and the standard deviation is smaller than the length of predetermined range.

5.2 Limitation and Future Work

The work presented here shows the feasibility of tuning power and connectivity of nodes in physical layer of WSN based on the path-loss. Additionally, the underlying statistics in short range path-loss modeling was investigated. There are some limitations which needs further research before the proposed algorithm can be applied to a real WSN.

In the case of path-loss slope estimation with MoM, here we assumed that path length of wireless links is uniformly, or normally distributed and standard deviation of distance error is given. Therefore, if path length in a path-loss dataset has other distributions, the MoM cannot be applied. Also, standard deviation of distance error in the case of NLOS communication is difficult to calculate and can only be captured with channel impulse response.

In the case of path-loss modeling, we showed the existence of long-term temporal and spatial variation of wireless links and how it could be modeled. Capturing path-loss of communication links with more than one sample in a day can explain the underlying reasons behind such behavior of wireless links. Also, analyzing a path-loss measurement dataset which capture the variation of wireless links for more than a year, would show any possible trends and patterns of long-term temporal variation in one year.

In the case of combined RNP-TPA algorithm for tuning WSN in physical layer, by integrating routing algorithm in network layer the performance of WSN in network layer before and after applying the combined algorithm can be assessed. Determining a standard approach to measure the irregularity of WSN and altering the RNP algorithm to only fill the relay nodes in the constrained locations are the other future works. Finally, use an LQI metric in physical layer beside path-loss can help increase the efficiency of the combined algorithm.

References

- [1] M. G. Kanabar, S. Member, I. Voloh, S. Member, and D. Mcginn, “Reviewing smart grid standards for protection , control , and monitoring applications,” in *Proc. IEEE PES*, 2011, pp. 1–8.
- [2] “IEEE guide for smart grid interoperability of energy technology and information technology operation with the electric power system (EPS), end-use applications, and loads,” in *Proc. IEEE Std 2030*, 2011, pp. 1–126.
- [3] V. Hamidi, K. S. Smith, and R. C. Wilson, “Smart grid technology review within the transmission and distribution sector,” in *Proc. IEEE PES*, 2010, pp. 1–8.
- [4] S. Supriya, M. Magheshwari, S. Sree Udhyalakshmi, R. Subhashini, and Musthafa, “Smart grid technologies: Communication technologies and standards,” *Int. J. Appl. Eng. Res.*, vol. 10, no. 20, pp. 529–539, 2015.
- [5] R. DeBlasio and C. Tom, “Standards for the smart grid,” in *2008 IEEE Energy 2030 Conference*, 2008, pp. 1–7.
- [6] R. Davies, “Hydro one’s smart meter initiative paves way for defining the smart grid of the future,” in *Proc. IEEE PES*, 2009, pp. 1–2.
- [7] W. Luan, D. Sharp, and S. Lancashire, “Smart grid communication network capacity planning for power utilities,” in *Proc. IEEE PES Transmission and Distribution Conference and Exposition*, 2010, pp. 1–4.
- [8] L. Kelly, A. Rowe, and P. Wild, “Analyzing the impacts of plug-in electric vehicles on distribution networks in British Columbia,” in *Proc. IEEE EPEC*, 2009, pp. 1–6.
- [9] R. E. Brown, “Impact of Smart Grid on Distribution System design,” in *Proc. IEEE PES*, 2008, pp. 1–4.
- [10] V. Dabic and D. Atanackovic, “Voltage VAR optimization real time closed loop deployment - BC Hydro challenges and opportunities,” in *Proc. IEEE PES*, 2015, pp. 1–5.
- [11] V. Keller *et al.*, “Electrification of road transportation with utility controlled charging: A case study for British Columbia with a 93% renewable electricity target,” *Appl. Energy*, vol. 253, 2019.
- [12] S. Bavarian, L. Lampe, C. Siew, S. Lancashire, and K. Adeleye, “Leveraging the smart metering infrastructure in distribution automation,” in *Proc. IEEE SmartGridComm*, 2012, pp. 157–162.
- [13] M. E. El-Hawary, “The smart grid - State-of-the-art and future trends,” *Electr. Power*

Components Syst., vol. 42, no. 3–4, pp. 239–250, 2014.

- [14] S. S. S. R. Depuru, L. Wang, and V. Devabhaktuni, “Smart meters for power grid: Challenges, issues, advantages and status,” *Renew. Sustain. Energy Rev.*, vol. 15, no. 6, pp. 2736–2742, 2011.
- [15] V. K. Sood, D. Fischer, J. M. Eklund, and T. Brown, “Developing a communication infrastructure for the smart grid,” in *Proc. IEEE EPEC*, 2009, pp. 1–7.
- [16] M. M. Rahman and A. Mto, “Technologies required for efficient operation of a smart meter network,” in *Proc. IEEE Conference on Industrial Electronics and Applications*, 2011, pp. 809–814.
- [17] “IEEE standard for local and metropolitan area networks--Part 15.4: low-rate wireless personal area networks (LR-WPANs) amendment 3: physical layer (PHY) specifications for low-data-rate, wireless, smart metering utility networks,” *IEEE Std 802.15.4g-2012 (Amendment to IEEE Std 802.15.4-2011)*, pp. 1–252, Apr. .
- [18] Kuor-Hsin Chang and B. Mason, “The IEEE 802.15.4g standard for smart metering utility networks,” in *2012 IEEE Third International Conference on Smart Grid Communications (SmartGridComm)*, 2012, pp. 476–480.
- [19] L. D. P. Mendes and J. J. P. C. Rodrigues, “A survey on cross-layer solutions for wireless sensor networks,” *J. Netw. Comput. Appl.*, vol. 34, no. 2, pp. 523–534, 2011.
- [20] T. Lee, X. Xie, and L. Chang, “RSSI-based IPv6 routing metrics for RPL in low-power and lossy networks,” in *Proc. IEEE SMC*, 2014, pp. 1714–1719.
- [21] H. R. Kermajani and C. Gomez, “Route change latency in low-power and lossy wireless networks using RPL and 6LoWPAN neighbor discovery,” in *Proc. IEEE ISCC*, 2011, pp. 937–942.
- [22] T. O. Olasupo, C. E. Otero, L. D. Otero, K. O. Olasupo, and I. Kostanic, “Path loss models for low-power, low-data rate sensor nodes for smart car parking systems,” *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 6, pp. 1774–1783, 2017.
- [23] S. Sun *et al.*, “Propagation path loss models for 5G urban micro-and macro-cellular scenarios,” in *Proc. IEEE VTC*, 2016, pp. 1–6.
- [24] D. W. Matolak, S. Member, Q. Zhang, and Q. Wu, “Path loss in an urban peer-to-peer channel for six public-safety frequency bands,” *IEEE Wirel. Commun. Lett.*, vol. 2, no. 3, pp. 263–266, 2013.
- [25] L. R. and S. K. S. Chandrasekharan, A. Al-Hourani, K. Magowe, “Propagation measurements for D2D in rural areas,” in *Proc. IEEE ICCW*, 2015, pp. 639–645.
- [26] M. C. Walden, T. Jackson, and W. H. Gibson, “Development of an empirical path-loss

- model for street-light telemetry at 868 and 915 MHz,” in *Proc. IEEE APSURSI*, 2011, no. 1, pp. 3389–3392.
- [27] S. Van Huffel, “Total least squares and errors- in-variables modeling: bridging the gap between statistics, computational mathematics and engineering,” in *Proc. COMPSTAT*, 2004, pp. 539–555.
- [28] A. Karttunen and K. Haneda, “Large-scale parameter estimation in channel sounding with limited dynamic range,” in *Proc. IEEE EuCAP*, 2019, pp. 1–5.
- [29] C. Gustafson, T. Abbas, D. Bolin, and F. Tufvesson, “Statistical modeling and estimation of censored pathloss data,” *IEEE Wirel. Commun. Lett.*, vol. 4, no. 5, pp. 569–572, 2015.
- [30] T. A. Jones, “Fitting straight lines when both variables are subject to error,” *J. Int. Assoc. Math. Geol.*, vol. 11, pp. 1–25, 1979.
- [31] C.-L. C. and J. W. Van Ness, “On estimating linear relationships when both variables are subject to errors,” *J. R. Stat. Soc. Ser. B (Statistical Methodol.)*, vol. 56, no. 1, pp. 167–183, 1994.
- [32] N. R. Draper, “Straight line regression when both variables are subject to error,” in *Proc. 3rd Annual Conference on Applied Statistics in Agriculture*, 1991, pp. 0–18.
- [33] A. G. Fragkaki, M. A. Koupparis, and C. G. Georgakopoulos, “Quantitative structure-retention relationship study of α -, β 1-, and β 2-agonists using multiple linear regression and partial least-squares procedures,” *Anal. Chim. Acta*, vol. 512, no. 1, pp. 165–171, 2004.
- [34] C. Y. Gau and M. A. Stadtherr, “Reliable nonlinear parameter estimation using interval analysis: Error-in-variable approach,” *Comput. Chem. Eng.*, vol. 24, no. 2–7, pp. 631–637, 2000.
- [35] G. E. . Box and D. R. Cox, “Regression models and life,” *J. R. Stat. Soc. Ser. B*, vol. 34, no. 2, pp. 187–202, 1972.
- [36] Y. Hu and S. Schennach, “Instrumental Variable Treatment of Nonclassical,” *Econometrica*, vol. 76, no. 1, pp. 195–216, 2008.
- [37] M. A. and K. R. A. Alsayyari, I. Kostanic, C. Otero, “An empirical path loss model for wireless sensor network deployment in a sand terrain environment,” in *Proc. IEEE WF-IoT*, 2014, pp. 218–223.
- [38] A. Aisayyari, I. Kostanic, and C. E. Otero, “An Empirical path loss model for wireless sensor network deployment in an artificial turf environment,” in *Proc. IEEE ICNSC*, 2014, pp. 637–642.
- [39] A. Alsayyari, I. Kostanic, and C. E. Otero, “An empirical path-loss model for wireless

- sensor network deployment in a concrete surface environment,” in *Proc. IEEE WAMICON*, 2015, pp. 1–6.
- [40] J. W. Gillard, “Error in variable regressions: What is the appropriate model?,” Cardiff university, 2007.
- [41] Y. Ming and L. Yi, “Model selection and estimation in regression with grouped variables,” *J. R. Stat. Soc. Ser. B (Statistical Methodol.)*, vol. 68, pp. 49–67, 2006.
- [42] J. W. Gillard and S. Road, “An historical overview of linear regression with errors in both variables by contents,” *Math. Sch. Cardiff Univ., Wales, UK, Tech. Rep*, 2006.
- [43] J. Gillard, “An overview of linear structural models in errors in variables regression,” *Revstat Stat. J.*, vol. 8, no. 1, pp. 57–80, 2010.
- [44] M. Hata, “Empirical formula for propagation loss in land mobile radio services,” *IEEE Trans. Veh. Technol.*, vol. 29, no. 3, pp. 1–9, 1980.
- [45] S. S. Ghassemzadeh, H. R. Worstell, and R. R. Miller, “Wireless neighborhood area network path loss characterization at 5 . 7 GHz,” in *Proc. IEEE VTC*, 2010, pp. 1–6.
- [46] H. Tsuchiya, “Characterization of the radio propagation channel in residential environment for smart meter communications,” in *Proc. IEEE Asia-Pacific Microwave Conference*, 2011, pp. 713–716.
- [47] V. Erceg *et al.*, “An empirically based path loss model for wireless channels in suburban environments,” *IEEE J. Sel. AREAS Commun.*, vol. 17, no. 7, pp. 1205–1211, 1999.
- [48] J. Turkka and M. Renfors, “Path loss measurements for a non-line-of-sight mobile-to-mobile environment,” in *Proc. IEEE ITST*, 2008, pp. 274–278.
- [49] T. S. Rappaport and others, *Wireless communications: principles and practice*, vol. 2. prentice hall PTR New Jersey, 1996.
- [50] H. Bai and M. Atiquzzaman, “Error modeling schemes for fading channels in wireless communications: A survey,” *IEEE Commun. Surv. Tutorials*, vol. 5, no. 2, pp. 2–9, 2003.
- [51] M. Younis and K. Akkaya, “Strategies and techniques for node placement in wireless sensor networks: A survey,” *Ad Hoc Networks*, vol. 6, no. 4, pp. 621–655, 2008.
- [52] I. Khoufi *et al.*, “Survey of deployment algorithms in wireless sensor networks : Coverage and connectivity issues and challenges,” *Int. J. Auton. Adapt. Commun. Syst.*, vol. 10, no. 4, pp. 341–390, 2014.
- [53] I. Khemapech, A. Miller, and I. Duncan, “A survey of transmission power control in wireless sensor networks,” in *Proc.PGNet*, 2007, pp. 15–20.

- [54] I. Khoufi, P. Minet, and A. Laouiti, "Fault-tolerant and constrained relay node placement in wireless sensor networks," in *Proc. MASS*, 2016, pp. 127–135.
- [55] W. Zhu, S. Xianhe, L. Cuicui, and C. Jianhui, "Relay Node Placement Algorithm Based on Grid in Wireless Sensor Network," in *Proc. International Conference on Instrumentation, Measurement, Computer, Communication and Control*, 2013, pp. 278–283.
- [56] A. Bari, A. Jaekel, J. Jiang, and Y. Xu, "Design of fault tolerant wireless sensor networks satisfying survivability and lifetime requirements," *Comput. Commun.*, vol. 35, no. 3, pp. 320–333, 2012.
- [57] J. Tang, B. Hao, and A. Sen, "Relay node placement in large scale wireless sensor networks," *Comput. Commun.*, vol. 29, no. 4, pp. 490–501, 2006.
- [58] D. Yang, S. Misra, X. Fang, G. Xue, and J. Zhang, "Two-tiered constrained relay node placement in wireless sensor networks: computational complexity and efficient approximations," *IEEE Trans. Mob. Comput.*, vol. 11, no. 8, pp. 1–9, 2012.
- [59] A. Chelli, M. Bagaa, D. Djenouri, I. Balasingham, and T. Taleb, "One-step approach for two-tiered constrained relay node placement in wireless sensor networks," *IEEE Wirel. Commun. Lett.*, vol. 5, no. 4, pp. 448–451, 2016.
- [60] A. Kashyap, S. Khuller, and M. Shayman, "Relay placement for higher order connectivity in wireless sensor networks," in *Proc. IEEE INFOCOM*, 2006, pp. 1–12.
- [61] C. Ma, W. Liang, M. Zheng, and H. Sharif, "A connectivity-aware approximation algorithm for relay node placement in wireless sensor networks," *IEEE Sens. J.*, vol. 16, no. 2, pp. 515–528, 2016.
- [62] S. Misra, S. D. Hong, G. Xue, and J. Tang, "Constrained relay node placement in wireless sensor networks to meet connectivity and survivability requirements," in *Proc. IEEE INFOCOM*, 2008, pp. 879–887.
- [63] S. Lee and M. Younis, "Optimized relay node placement for connecting disjoint wireless sensor networks," *Comput. Networks*, vol. 56, no. 12, pp. 2788–2804, 2012.
- [64] F. Senel and M. Younis, "Relay node placement in structurally damaged wireless sensor networks via triangular steiner tree approximation," *Comput. Commun.*, vol. 34, no. 16, pp. 1932–1941, 2011.
- [65] K. Jain, J. Padhye, V. N. Padmanabhan, and L. Qiu, "Impact of interference on multi-hop wireless network performance," *Wirel. Networks*, vol. 11, no. 4, pp. 471–487, 2003.
- [66] R. Ramanathan and R. Rosales-Hain, "Topology control of multihop wireless networks using transmit power adjustment," in *Proc. IEEE INFOCOMM*, 2000, pp. 404–413.

- [67] M. Kubisch, H. Karl, A. Wolisz, L. C. C. Zhong, and J. Rabaey, "Distributed algorithms for transmission power control in wireless sensor networks," in *Proc. WCNC*, 2003, pp. 558–563.
- [68] D. Son, K. Bhaskar, and H. John, "Experimental study of the effects of transmission power control and blacklisting in wireless sensor networks," in *Proc. IEEE SECON*, 2004, pp. 289–298.
- [69] J. Jeong, D. Culler, and J. H. Oh, "Empirical analysis of transmission power control algorithms for wireless sensor networks," in *Proc. IEEE INSS*, 2007, pp. 27–34.
- [70] G. Liu, Z. Li, X. Zhou, and S. Li, "ATPC: Adaptive Transmission Power Control for wireless sensor networks," *ACM Trans. Sens. Networks*, vol. 12, no. 1, pp. 1–33, 2016.
- [71] J. Kim, S. Chang, and Y. Kwon, "ODTPC: On-demand Transmission Power Control for Wireless Sensor Networks," in *Proc. IEEE ICOIN*, 2008, pp. 1–5.
- [72] S. T. Cheng and M. Wu, "Optimization of multilevel power adjustment in wireless sensor networks," *Telecommun. Syst.*, vol. 42, no. 1–2, pp. 109–121, 2009.
- [73] C. Zhu, C. Zheng, L. Shu, and G. Han, "A survey on coverage and connectivity issues in wireless sensor networks," *Journal of Network and Computer Applications*, vol. 35, no. 2. Elsevier, pp. 619–632, 2012.
- [74] Y. Wang, "Topology Control for Wireless Sensor Networks," *Wirel. Sens. Networks Appl.*, pp. 113–147, 2008.
- [75] C. S. Sum, F. Kojima, and H. Harada, "Coexistence of homogeneous and heterogeneous systems for IEEE 802.15.4g smart utility networks," in *Proc. DySPAN 2011*, 2011, pp. 510–520.
- [76] D. Li and H. Liu, "The effect of deployment irregularity on coverage in wireless sensor networks," in *proc. ISSNIP*, 2005.
- [77] IEEE Standard, "IEEE standard for local and metropolitan area networks. Part 15.4 : Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications," *IEEE Std 802.15.4*, pp. 1–26, 2009.
- [78] C. Bettstetter, "On the minimum node degree and connectivity of a wireless multihop network," in *Proc. of the International Symposium on Mobile Ad Hoc Networking and Computing*, 2002, pp. 80–91.
- [79] M. D. Penrose, "On k-connectivity for a geometric random graph," *Random Struct. Algorithms*, vol. 15, no. 2, pp. 145–164, 1999.
- [80] D. G. Kirkpatrick and R. Seidel, "The Ultimate Planar Convex Hull Algorithm?," *SIAM J. Comput.*, vol. 15, no. 1, pp. 287–299, 1986.

- [81] R. L. Graham, “An efficient algorithm for determining the convex hull of a finite planar set,” *Inf. Process. Lett.*, vol. 1, no. 4, pp. 132–133, 1972.
- [82] S. J. L. Lancashire, “Modelling short-range path loss in smart meter mesh networks,” University of British Columbia, 2016.