

Analysis of Wireless Multi-Hop Broadcasting: Optimal and Approximate Solutions and Applications

by

Mohammadreza Ataei Naeini

A dissertation submitted to the
Faculty of Graduate and Postdoctoral Affairs
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Electrical and Computer Engineering

Ottawa-Carleton Institute for Electrical and Computer Engineering
Department of Systems and Computer Engineering
Carleton University
Ottawa, Ontario
July, 2015

©Copyright
Mohammadreza Ataei Naeini, 2015

The undersigned hereby recommends to the
Faculty of Graduate and Postdoctoral Affairs
acceptance of the dissertation

Analysis of Wireless Multi-Hop Broadcasting: Optimal and Approximate Solutions and Applications

submitted by **Mohammadreza Ataei Naeini**

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Electrical and Computer Engineering

Professor Amir H. Banihashemi, Thesis Co-Supervisor

Professor Thomas Kunz, Thesis Co-Supervisor

Professor Shahrokh Valaee, External Examiner

Professor Roshdy Hafez, Chair,
Department of Systems and Computer Engineering

Ottawa-Carleton Institute for Electrical and Computer Engineering
Department of Systems and Computer Engineering
Carleton University
July, 2015

Abstract

In this work, we design and analyze transmission range assignments for broadcasting in wireless multi-hop networks. Moreover, we study different features of wireless networks. We consider network scenarios in which the exact location of the nodes is known and others where the nodes location is known probabilistically. For the former, we propose optimal and near-optimal algorithms to solve the *Minimum-Energy Broadcasting* problem for linear (one-dimensional) networks. We further extend our solutions to encompass cross networks, in which the nodes are located on two perpendicular lines. The proposed algorithms have polynomial-time complexity, and are shown to perform better than previously known algorithms (for some cases, they are the first polynomial-time solutions).

For probabilistic networks, we propose a transmission range assignment such that for a given average total consumed power, the linear network is connected with high probability. We then analyze some features of these networks, including derivation of exact formulas for the probability of connectivity of any location of the network to the source, the hop-count probability mass function (pmf) of an arbitrary location of the network, and the pdf of the maximum coverage (last reachable distance from the source) for a given number of hops. The proposed analyses are applicable to networks with non-identical transmission range assignments, where the nodes are placed independently and identically according to a Poisson distribution with an arbitrary density function.

Based on the derived formulas, we then propose localization and location verification methods. We show that our proposed localization method not only has a competitive performance for a range-free method, but also outperforms range-based methods with a local distance measurement error of 10% or more. Furthermore, the proposed location verification protocol is shown to have better results compared to the existing verification systems that also use the hop-count information. We also evaluate the proposed schemes in the presence of Rician fading and show that their performance is rather robust with respect to the change in the fading parameter.

I dedicate this dissertation to my wonderful family. Particularly to my late beloved brother, Mehdi, whose influence on my life will be with me forever. May he rest in peace while his love and memory is always in my heart.

Acknowledgments

First of all, I would like to express my deepest gratitude to my supervisors Dr. Amir H. Banihashemi, and Dr. Thomas Kunz for their excellent guidance and advices. Without their patience and continuous and tireless support and assistance, the completion of this thesis would not have been possible. Finally, I would like to extend my deepest thanks to those whose support cannot be forgotten, my family and especially my wife Azadeh, for their endless support in all my academic and personal endeavors.

Table of Contents

Abstract	iii
Acknowledgments	v
Table of Contents	vi
List of Tables	ix
List of Figures	x
Nomenclature	xiii
1 Introduction	1
1.1 Motivation	1
1.2 Our Contributions	3
1.3 Published and Submitted Work	6
1.3.1 Published Journal Papers	6
1.3.2 Submitted Journal Paper	7
1.3.3 Published Conference Papers	7
1.4 Organization of the Thesis	7
2 Previous Works	8
2.1 Energy-Efficient Broadcasting	8
2.2 Probabilistic Networks	10
2.3 Broadcasting Features Analysis	12
3 Energy-Efficient Broadcasting for Linear Networks	16
3.1 System Model	17
3.2 Energy-Efficient Range Assignments	18

3.2.1	Near-Optimal Range Assignment with Linear Time Complexity	19
3.2.2	Optimal Range Assignment	22
3.2.3	Distributed Range Assignment	26
3.2.4	Identical Transmission Range	27
3.3	Numerical Results	28
4	Energy-Efficient Broadcasting for Cross Networks	33
4.1	System Model	33
4.2	Proposed Range Assignments	37
4.2.1	Optimal Range Assignment	37
4.2.2	Near-Optimal Range Assignment With Linear Time Complexity	56
4.2.3	Distributed Range Assignment	58
4.3	Special Case: Source at the Intersection	61
4.4	A More General Case: Grid Networks	61
4.5	Asymptotic Analysis	62
4.5.1	Step 1:	63
4.5.2	Step 2:	67
4.6	Numerical Results	69
5	Probabilistic Linear Networks: Range Assignment	73
5.1	System Model	74
5.2	Transmission Range Assignment	76
5.3	Probability of Connectivity	79
5.4	Contention Analysis	79
5.5	Numerical Results	81
6	Probabilistic Linear Networks: Features Analysis	91
6.1	System Model	92
6.1.1	Continuous Network Model	92
6.1.2	Discrete Network Model	93
6.2	Probability of Connectivity Analysis	95
6.2.1	Probability of Connectivity for the Continuous Network Model	96
6.2.2	Probability of Connectivity for the Discrete Network Model .	98
6.3	Hop-Count Analysis	99
6.3.1	Bounds	100

6.3.2	Hop-Count Analysis for the Continuous Network Model	102
6.3.3	Hop-Count Analysis for the Discrete Network Model	104
6.4	Last Reachable Distance Analysis	106
6.4.1	Last Reachable Distance Analysis for the Continuous Network Model	106
6.4.2	Last Reachable Distance Analysis for the Discrete Network Model	108
6.4.3	Obtaining the Hop-Count pmf using the Last Reachable Distance Information	110
6.5	Numerical Results	111
6.5.1	Model Used for Simulations	111
6.5.2	Results	112
7	Localization and Location Verification	119
7.1	System Model	120
7.2	Hop-Count Analysis for Localization and Location Verification	120
7.3	Localization	122
7.3.1	Enhancement	125
7.4	Location Verification	128
7.4.1	Plausibility Calculation	128
7.4.2	Performance Analysis	130
7.5	Numerical Results	132
7.5.1	Localization	134
7.5.2	Location Verification	138
8	Conclusion and Future Work	141
8.1	Conclusion	141
8.2	Future Work	142
8.2.1	Average Probability of Connectivity	143
List of References		148

List of Tables

4.1	The simulation results of cross networks with source at intersection.	70
4.2	The simulation results of general cross networks.	70
7.1	Hop-count pmf of the location $x = 450$ meters, receiving the message from the anchor at $x = 0$	133
7.2	Hop-count pmf of the verifier at $x = 0$, receiving the message from a node at location $x = 500$ meters.	133
7.3	Some elements of the localization table \mathfrak{T} with $N_{h_1,h_2} = 3$, for $\gamma = 1$. . .	138
7.4	Some elements of the localization table \mathfrak{T} for different values of N_{h_1,h_2} , (for $h_1 = 5$, $h_2 = 6$ and $\gamma = 1$).	139

List of Figures

1.1	(a) Network realization in which $h(x) = 4$. (b) Network realization in which $h(x) = 5$	5
3.1	In an energy-efficient range assignment, some nodes do not need to transmit.	20
3.2	Comparison of the energy consumption of different range assignments.	29
3.3	The histogram of the normalized differences between the energy consumptions of the proposed optimal, linear near-optimal and distributed algorithms for $\lambda = 0.03$	30
3.4	Network topologies where the difference between the consumed energy of different algorithms (linear near-optimal vs. distributed in (a), and optimal vs. linear suboptimal in (b)) can be large.	31
3.5	The histogram of the distance between node b_m and the source node for $\lambda = 0.03$	32
4.1	Some definitions/notations for networks on a cross.	35
4.2	Segmentation of cross networks.	35
4.3	The next adjacent neighbors of nodes a, b, c, d and e	36
4.4	The graph G^\diamond , for which its MST is needed to be found in each of the nodes in the set $\{l_{\text{II}}, f_{\text{III}}, f_{\text{IV}}, f_{\text{V}}\}$	58
4.5	A possible structure for grid networks with perpendicular line-segments.	61
4.6	Energy consumption comparison of different range assignments.	71
4.7	Comparison of the energy consumption of different range assignments for two by two square gird networks.	72
5.1	Network model and the proposed transmission range assignment.	77
5.2	Density function $\lambda(x)$ along with the range assignment of the proposed scheme for $\mathcal{K} = 4$ and 6 , in the first scenario.	83

5.3	Comparison of the probability of connectivity of the network with density function $\lambda(x)$ given in Fig. 5.2 for the two cases of the proposed range assignment $R(x)$ and the identical transmission range assignment $R_{identical}$ for $\mathcal{K} = 4$ and $\mathcal{K} = 6$.	84
5.4	Density function given in [1] as a real-world example along with the range assignment of the proposed scheme for $\mathcal{K} = 6$.	85
5.5	Comparison of the probability of connectivity for a real-world example with density function $\lambda(x)$ given in Fig. 5.4 for the two cases of the proposed range assignment $R(x)$ and the identical transmission range assignment $R_{identical}$ ($\mathcal{K} = 6$).	86
5.6	Comparison of the average contention levels for the network with the density function $\lambda(x)$ given in Fig. 5.2 for the two cases of the proposed range assignment $R(x)$ and the identical transmission range assignment $R_{identical}$ ($\mathcal{K} = 4$ and $\mathcal{K} = 6$).	87
5.7	Comparison of the average contention levels of the network with the density function $\lambda(x)$ given in Fig. 5.4 for the two cases of the proposed range assignment $R(x)$ and the identical transmission range assignment $R_{identical}$ ($\mathcal{K} = 6$).	88
5.8	Density function $\lambda(x)$ along with the range assignment of the proposed scheme for $\mathcal{K} = 4$ and 6, in the third scenario.	89
5.9	Comparison of the probability of connectivity of the network with density function $\lambda(x)$ given in Fig. 5.8 for the two cases of the proposed range assignment $R(x)$ and identical transmission range assignment $R_{identical}$ for $\mathcal{K} = 4$ and $\mathcal{K} = 6$.	89
5.10	Comparison of the average contention levels for the network with density function $\lambda(x)$ given in Fig. 5.8 for the two cases of the proposed range assignment $R(x)$ and identical transmission range assignment $R_{identical}$ for $\mathcal{K} = 4$ and $\mathcal{K} = 6$.	90
6.1	A network configuration, to show that the interval $(d^{min}[h-2], d^{max}[h-1])$ must be searched to find $d^{min}[h]$.	101
6.2	Probability of connectivity of the network with the given density function in Fig. 5.4 obtained by our analysis and simulation for the identical and the proposed transmission range assignments ($\mathcal{K} = 6$).	113

6.3	Hop count pmf of segment 4,000 (400 meters away from the source), obtained by analysis and simulation (for the proposed transmission range assignment).	114
6.4	Hop count pmf of segment 6,000 (600 meters away from the source), obtained by analysis and simulation (for the proposed transmission range assignment).	115
6.5	The data reception probability up to a certain number of hops for the identical transmission range assignment.	116
6.6	The data reception probability up to a certain number of hops for the proposed transmission range assignment.	117
6.7	Probability for the last reachable location by at most h hops (h values are indicated in the figure) be greater than or equal to the segment on the x -axis (using Equation (6.41)).	118
7.1	Detection happens for $\check{x} < x^* - d$ (interval Δ_L^D) and $\check{x} > x_\tau^R$ (interval Δ_R^D). False Alarm happens for $x^* - d < \check{x} < x_\tau^L$ (interval Δ^F).	131
7.2	The pdf of the location of a node that receives the message by $h = 5$ hops from the anchor at $x = 0$	134
7.3	The pdf of the location of a node that receives the message by $h_1 = 5$ and $h_2 = 6$ hops from two anchors at $x = 0$ and $x = L = 700$ meters, respectively.	135
7.4	Localization error of different approaches for different densities, while $\forall x : R_a^{(1)}(x) = R_a^{(2)}(x) = 70$ meters.	137
7.5	Plausibility of all the locations in the interval [300, 450] meters for different approaches, where the claim is from a node located at $x = 375$ meters.	139
7.6	Receiver operating characteristic (ROC) curve of different approaches.	140
8.1	The identical transmission range assignment along with three other heuristic assignments (all the four of these assignments consume the same amount of power on average).	145
8.2	Probability of connectivity of a network with uniform node distribution employing the transmission range assignments shown in Fig. 8.1.	146
8.3	Average probability of connectivity up to the location of the x-axis in a network with uniform node distribution employing the transmission range assignments shown in Fig. 8.1.	147

Nomenclature

ACC	Adaptive Cruise Control
a.a.s.	Asymptotically Almost Surely
BIP	Broadcast Incremental Power
CCA	Curvilinear Component Analysis
DV	Distance Vector
FFT	Fast Fourier Transform
MAC	Multiple Access Control
MAP	Maximum A Posterior
MCDS	Minimum Connected Dominating Set
MDS	Multi-Dimensional Scaling
MMSE	Minimum Mean-Square Error
MST	Minimum Spanning Tree
NP	Non-deterministic Polynomial-time
PLV	Probabilistic Location Verification
pdf	Probability Density Function
pmf	Probability Mass Function

QoS	Quality of Service
RND	Regulated Neighborhood Distance
ROC	Receiver Operating Characteristic
RSSI	Received signal Strength Indication
SPT	Shortest-Path Tree
TDOA	Time Difference of Arrival
TOA	Time of Arrival
VANET	Vehicular Ad-Hoc Network

Chapter 1

Introduction

1.1 Motivation

Wireless ad-hoc networks have attracted more interest in recent years due to their numerous applications [2–4]. In these networks, broadcasting mechanisms are used for data exchange purposes, e.g., disseminating important messages or path discovery information in routing algorithms [5]. Studying different features of wireless networks such as reliability, energy-efficiency and delay, helps in analyzing and designing such networks, e.g., to guarantee an acceptable level of quality of service (QoS). Reliability can be translated to the probability of connectivity of the network. Delay and energy-consumption are related to the multi-hop structure and transmission range assignment of the network [6–8].

In this research, we first approached the *Minimum-Energy Broadcasting* problem. This problem focuses on finding a transmission range assignment for all the nodes in the network such that the total consumed energy for broadcasting data from one specific node (a.k.a the source node) to all the other nodes is minimized [9]. A transmission range assignment defines the range each network node has to transmit the data with. In this problem, the exact location of the nodes is known¹.

We further extended our research to include cases where the knowledge of the exact location of the nodes is not available, and the locations are known probabilistically. For the probabilistic model, we proposed a transmission range assignment such that

¹In actual wireless ad-hoc networks (except for mesh networks) the network topology can be assumed to be fixed in finite, periodical time-frames. One has to solve the minimum-energy broadcasting problem for each time-frame. This results in the need to obtain low time complexity algorithms for finding the solution for that problem.

for a given average total consumed power, the network is connected with high probability. Furthermore, we derived exact formulas for the probability of connectivity of any location of the network to the source, the hop-count² probability mass function (pmf) of an arbitrary location of the network, and the probability density function (pdf) of the maximum coverage (last reachable distance from the source) for a given number of hops. Using the techniques described in [7] and [8] (as just two examples), the derived formulas can be used to estimate the delay that the data experiences to be delivered to all the nodes within a desired coverage range. Moreover, as we show later, this information is used to localize nodes (i.e., estimating the location of the nodes), and verify location claims made by nodes in an ad-hoc network.

In wireless ad-hoc networks, location information of network nodes is essential to support location-based services. Some examples of location-based services are: tracking events and targets (especially in wireless sensor networks), supporting geographic routing protocols, and providing enhanced security protection mechanisms. Localization protocols have been proposed to derive a node's location without the use of localization hardware (such as GPS receivers), and thus to reduce the node's cost. Localization enables signal processing algorithms to make inference in the spatial domain about the environment in which they are placed [10]. For an extensive survey on localization algorithms see [11, 12]. Furthermore, localization is a key element in location verification algorithms [13, 14]. Location verification mechanisms use the location information to verify a claimed location [15]. These mechanisms are used to confront attacks which take advantage of a lack of location information. For example, malicious nodes can advertise false positions to alter the routing paths, etc. Moreover, location verification enables location-based access control [16], where the users with verified locations can access particular resources according to a given policy.

In the most part of our research, we considered one-dimensional wireless ad-hoc networks where the nodes are located on a line. Linear networks are used, for example, to model the wireless communication in a vehicular ad hoc network (VANET), where the data transmission is along a road [17], or to model sensor networks, where the sensors are distributed in a one-dimensional sensor field [18]. An example of a linear VANET is a platoon of vehicles, where all the vehicles follow a leading vehicle (for example, a truck). Each vehicle is equipped with adaptive cruise control (ACC), to maintain its speed while using radar to keep a safe distance from the vehicle in

²For any location in the network, the hop-count is considered as the number of hops needed for the data to travel from the source to that specific location.

front of it [19]. In such a scenario, the leader broadcasts the information through a wireless channel to all the vehicles, in order to inform them of any turns or braking actions [20]. For an extensive survey on different communication protocols in VANETs see [21]. An example of a linear sensor network would be one to monitor and maintain power grids for ensuring high quality and reliability of power supply and also to monitor pipelines [22, 23]. We addressed the minimum-energy broadcasting in cross networks as well. Cross networks can be used in VANETs and sensor networks used for monitoring power grids and pipelines. Further extension of the minimum-energy broadcasting problem is known to be non-deterministic polynomial-time (NP) hard (for D -dimensional spaces with $D \geq 2$), [6, 24, 25].

It should be noted that the effort to make the communication system consume less energy (i.e., be more energy-efficient) is commonly understandable for sensor networks (as they run on battery), while there are some controversies for applying it to VANETs. The reason for those disagreements is that vehicles with a combustion engine are often considered to have an infinite amount of energy as the battery is constantly being charged while the engine runs. The reality is that with moving more towards electric vehicles, this will not be the case anymore. Also, one should always try to enhance the energy consumption of any system, to be less harmful to the environment (being more green), and to save energy for other tasks that need energy, and can perform better with higher energy budget.

1.2 Our Contributions

In this thesis, we first proposed transmission range assignments to reduce the energy consumption of wireless ad-hoc networks. The consumed energy of a node is assumed to be a direct function of the transmission range of that node. There are some other aspects to the energy consumption of a wireless device, such as the used coding, modulation, etc., but this research just focuses on the effect of transmission range of the nodes on energy consumption.

We proposed two centralized and one distributed solution for the minimum-energy broadcasting problem in linear networks. The proposed optimal solution finds the exact transmission range assignment for minimum-energy broadcasting, and has time complexity $\mathcal{O}(N^2)$, where N is the number of network nodes. This improves the time

complexity $\mathcal{O}(N^3)$ of existing solutions. The other centralized algorithm is linear-time and finds an approximation of the optimal solution. Furthermore, we proposed a simple distributed range assignment algorithm for energy-efficient broadcasting. We demonstrated that on average both the linear-time approximation and the distributed algorithm are almost as efficient as the optimal range assignment for networks with uniformly distributed nodes. The distributed algorithm would be of particular interest not only because of its distributed nature, but also for its very low time complexity (constant in network size), and the small amount of network knowledge that each node requires to perform the algorithm (only the distances to the adjacent neighbors).

In addition to solving the minimum-energy broadcasting problem for linear networks, we obtained the optimal and some sub-optimal solutions for this problem for the networks located on a cross. Our solutions consist of an algorithm which finds the optimal assignment, and a near-optimal algorithm with close to optimal energy consumption. Moreover, we present a distributed algorithm that gives acceptable results. We show that the optimal solution for cross networks can be found in polynomial time, but has high time complexity ($\mathcal{O}(N^8)$ or $\mathcal{O}(N^{12})$). Our proposed near-optimal algorithm achieves near optimal results, and the proposed distributed algorithm performs close to the BIP (and BIP with sweep) algorithm for this kind of networks. The proposed algorithms have very low time complexity ($\mathcal{O}(N)$ for the centralized near-optimal algorithm and $\mathcal{O}(1)$ for the distributed algorithm). Furthermore, we proved that our distributed algorithm has the same outcome as the MST algorithm for networks on a cross, but with much lower time complexity ($\mathcal{O}(1)$ instead of $\mathcal{O}(N^2)$).

We also showed that the proposed distributed algorithm can be used for more general two-dimensional networks, where the nodes are located on a grid. The grid is not necessarily a square grid, but it has to have perpendicular line-segments. For these networks again, the proposed distributed algorithm is shown to perform close to the BIP (and BIP with sweep) algorithm, with much less time complexity and in a distributed manner. In addition to the consumed energy of the assignments, we also studied the number of hops needed for each node to receive the data for the first time (a.k.a., the hop-count of that node) for different assignments, and compare them.

To generalize our study, we considered wireless ad hoc networks in which the nodes are distributed according to a non-homogeneous Poisson distribution with an arbitrary density function. It should be noted that a network with uniform node distribution is

a special case of the networks discussed here (with constant density Poisson distribution). For this kind of networks, we proposed a transmission range assignment. For a given average power consumption, the proposed assignment can achieve a higher probability of connectivity compared with an identical transmission range assignment commonly used in the literature, e.g., in [1]. Furthermore, we showed that the proposed assignment, on average, has bounded contention levels³, while the contention level for the identical assignment can take much larger values.

We analyzed some major features of linear wireless ad hoc networks with Poisson spatial distribution of nodes with arbitrary density function. We derived exact formulas for the probability of connectivity of any location of the network to the source, the hop-count pmf of an arbitrary location of the network, and the pdf of the maximum coverage (last reachable distance from the source) for a given number of hops. Fig. 1.1 illustrates two network realizations along with the transmission ranges of the nodes. In our study, the transmission ranges of the nodes can be non-identical. In the network realization shown in Fig. 1.1(a), the hop-count of x , $h(x)$, needed for the data to be transmitted from the source node to location x is 4. We have $h(x) = 5$ in the network realization shown in Fig. 1.1(b). We derive the pmf of $h(x)$ for all the locations $0 < x \leq L$ of a linear network with length L . Furthermore, we derive the pmf of the number of hops needed for a node located at one end of the network to receive a message from a node at an arbitrary location of the network.

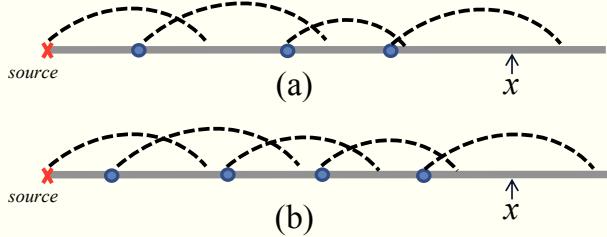


Figure 1.1: (a) Network realization in which $h(x) = 4$. (b) Network realization in which $h(x) = 5$.

Using the obtained pmf of $h(x)$, we propose a localization protocol that provides the pdf of the location of a node, given that it receives the message with a certain number of hops. Furthermore, we use our hop-count analysis to verify a location claim by calculating its plausibility.

³Contention level is defined as the number of simultaneous signal receptions in any location of the network.

This study, to the best of our knowledge, is the first to derive *exact* formulas for the analysis of wireless networks, as described above. In [26], the analysis is based on Gaussian approximation and in [27], it is based on a recursive formulation that does not lend itself to a non-recursive closed form. Moreover, both [26] and [27] perform the analysis only for the identical transmission range assignment. In our work, however, the range assignment is more general.

We show that our analysis obtains better localization results compared with [28–30], and even better than range-based methods when their local distance measurement has a typical error (i.e., more than 10%). Furthermore, our analysis results in improving the outcome of [13] for location verification, as we provide exact, and not approximate, formulas for the pmf of the hop-count random variables. This work, to the best of our knowledge, is the first to derive *exact* formulas for the hop-count analysis of wireless networks, as described above.

In addition to localization and location verification, our hop-count analysis can be used in studying efficiency of broadcasting mechanisms in wireless networks. Studying different features of wireless networks, such as energy-efficiency and delay, helps in analyzing and designing such networks, e.g., to guarantee an acceptable level of quality of service (QoS).

1.3 Published and Submitted Work

The following list of published and submitted works summarizes the peer-reviewed papers that appeared as a result of this work.

1.3.1 Published Journal Papers

- M. R. Ataei, T. Kunz, A. H. Banihashemi, "Localization and location verification in non-homogeneous one-dimensional wireless ad-hoc networks," *IEEE Journal on Selected Areas in Communications, Special Issue on Location-Awareness for Radios and Networks*, vol. 33, no. 7, May 2015.
- M. R. Ataei, A. H. Banihashemi, T. Kunz, "An energy-efficient transmission range assignment for 1-D wireless networks," *IEEE Wireless Communications Letters*, vol. 2, no. 5, pp. 543–546, October 2013.

- M. R. Ataei, A. H. Banihashemi, T. Kunz, "Low-complexity energy-efficient broadcasting in one-dimensional wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 61, pp. 3276-3282, September 2012.

1.3.2 Submitted Journal Paper

- M. R. Ataei, A. H. Banihashemi, T. Kunz, "Energy-Efficient Broadcasting for Cross Wireless Ad-Hoc Networks," *Submitted to IEEE Transactions on Wireless Communications*.

1.3.3 Published Conference Papers

- M. R. Ataei, A. H. Banihashemi, T. Kunz, "Minimum-energy broadcasting for cross wireless ad-hoc networks," *Proceedings of IEEE ICC 2015*, London, UK, June 8-12, 2015.
- M. R. Ataei, T. Kunz, A. H. Banihashemi, "Localization in non-homogeneous one-dimensional wireless ad-hoc networks," *Proceedings of IEEE ICC 2015*, London, UK, June 8-12, 2015.

1.4 Organization of the Thesis

The rest of the thesis is organized as follows. The previous researches done on the subjects of this thesis are explained in Chapter 2. In Chapter 3, we discuss our solutions for the minimum-energy broadcasting problem in linear networks. The proposed solutions for this problem in cross networks is presented in Chapter 4. We move on to the networks with probabilistic nodes locations, by first proposing a transmission range assignment in Chapter 5 to attain high probability of connectivity. The other features of this kind of networks are studied in Chapter 6. In Chapter 7 we demonstrate two applications of the obtained analysis, which are localization and location verification. A summary of what we have done so far, along with some suggested future research are presented in Chapter 8.

Chapter 2

Previous Works

2.1 Energy-Efficient Broadcasting

Given a specific source node that broadcasts the data to all the other nodes of the network, the problem of determining a set of retransmitting nodes and their corresponding transmission ranges, such that the sum of consumed node energies is minimized, is known as the minimum-energy broadcast problem. This problem is known to be NP hard for D -dimensional spaces with $D \geq 2$, [6, 24, 25]. One property that makes this problem hard is the *wireless multicast advantage* [31], i.e., reception of the transmitted data by multiple nodes within the range of a single transmission. There have been many works on solving this problem. For an extensive survey on the existing works on minimum-energy broadcasting (as a special case of multicasting) one can see [32].

For linear networks, the authors of [33] solved the problem of finding a range assignment for N nodes in a linear network for broadcasting from a specific node to all the other nodes with minimum energy. In [33], the nodes locations on the line are arbitrary and the source node is assumed to be known. It is also assumed that the network topology is known and available for solving the problem. The algorithm proposed in [33] has time complexity $\mathcal{O}(N^3)$ ¹. More recently, there have been some studies [34, 35] focusing on the same problem with additional constraints. In [34], the condition that all the nodes have to receive data in at most h hops has been added, and the proposed algorithm has time complexity $\mathcal{O}(hN^2)$. In [35], the same problem is solved with the assumption that the consumed energy in each node depends not

¹By definition, $f(x) = \mathcal{O}(g(x))$ if and only if there exists a positive constant a such that for all sufficiently large values of x , we have $|f(x)| \leq a|g(x)|$.

only on the transmission range of the node but also on an arbitrary positive weight assigned to the node. The problem is solved with algorithms of time complexity $\mathcal{O}(N^3)$ for the unconstrained case (i.e., $h = N - 1$) and $\mathcal{O}(hN^4)$ for the h -hop constrained case.

Some studies aimed to solve the minimum-energy broadcasting problem heuristically for general D -dimensional (with $D \geq 1$) networks. The broadcast incremental power (BIP) algorithm, with time complexity $\mathcal{O}(N^3)$ [31, 36], is one of the most commonly used methods for the energy-efficient broadcasting problem in 2-D networks. In [37], it is shown that the approximation ratio² of BIP, unlike some other well-known algorithms, e.g., shortest-path tree (SPT) [38], is a constant (does not depend on the total number of the nodes in the network). This constant is superior to the approximation ratio of the other algorithms also studied in [37]. In BIP, the goal is to construct a Minimum Spanning Tree (MST) of the network graph³ starting from the source, and adding new nodes one by one to the tree. The cost function, which is to be minimized in the construction, is the incremental cost of adding each additional node. This incremental cost is defined as the minimum additional power required of some node in the current tree to reach the new node. Recently in [39], the time complexity and the lower-bound of the approximation ratio of the BIP algorithm are strengthened. According to [39] the BIP algorithm has time complexity $\mathcal{O}(N^2)$.

A modification to BIP, called *sweep*, is proposed in [31, 36], to improve the power consumption. In this procedure, unnecessary transmissions are eliminated. It is shown in [40] that the time complexity of the sweep procedure is also $\mathcal{O}(N^2)$.

Another frequently cited method for energy-efficient broadcasting is the algorithm based on finding the Euclidean MST of the graph representing the network rooted at the source node [41]. The network graph is constructed by considering the nodes of the network as vertices of the graph. There is an edge between any two nodes, with weight equal to the Euclidean distance between the nodes. In this algorithm, the MST of the network graph is found (e.g., using Prim's algorithm). In the MST-based range assignment, for each node, the maximum weight of the edge between that node and its children⁴ is assigned as the node's transmission range, i.e.,

²The ratio of the energy consumption of a given assignment to that of the optimal assignment is called the approximation ratio of that assignment.

³The MST of a graph is the spanning tree with weight less than or equal to the weight of every other spanning tree of that graph.

⁴For any node in a tree, the nodes directly below it in the tree hierarchy are called the children of that node.

$$R^{MST}(i) = \max_{u:(i,u) \in MST} \{d(i, u)\}, \quad (2.1)$$

where $R^{MST}(i)$ denotes the assigned range to node i , and the notation (i, u) is used to denote the edge between nodes i and u in the MST, where i is the parent of u . The Euclidean distance between nodes a and b is denoted by $d(a, b)$.

The MST-based algorithm has time complexity $\mathcal{O}(N^2)$ [41]. As the nodes are mostly placed randomly in the network, the assumption of having unique weights for the edges of the corresponding network graph is commonly used. This assumption results in having a unique MST for the graph with probability one. Therefore the MST-based range assignment is unique. In [42, 43] it is shown that even for 2-D networks with special topologies, i.e., when the nodes are located at the intersection points of a square grid, the MST-based range assignment is far from optimal. It is worth mentioning that the MST-based range assignment is the optimal solution for the minimum-energy broadcasting problem in wired networks. The performance of BIP is shown to be better than that of the MST algorithm for general 2-D wireless networks [31, 36].

Papers [25], [6], and [44] studied the range assignment when the network graph is connected, i.e., each network node is connected to every other node via some path. This assignment is called a *complete range assignment*. The study in [25] focused on D -dimensional networks with $D \geq 2$, where every path between any two nodes consists of at most h hops. In [6], Kirousis *et al.* presented a dynamic programming algorithm with $\mathcal{O}(N^4)$ time complexity for finding a minimum cost complete range assignment, where the nodes are located on a line and the distances between the nodes are arbitrary. In [44], a minimum cost complete range assignment solution was found for linear networks with an algorithm with time complexity $\mathcal{O}(N^3)$, improving the time complexity of the algorithm of [6] by a factor of N .

2.2 Probabilistic Networks

For the case where only the node distribution along the network is known, rather than the exact location of the nodes, the goal usually is to determine the transmission range assignment such that the network is connected with high probability. In

most cases, the problem is studied under the assumption that the nodes are uniformly and independently distributed along the line, and that all the nodes have the same transmission range [45], [46]. In [45], it is shown that for a linear network of length L , where the density of the uniformly distributed nodes is λ , the network will be connected with probability one as the length of the network goes to infinity, if all the nodes have an identical transmission range⁵ r greater than $\frac{\ln(\lambda L)}{\lambda}$. Santi *et al.* [46] provided both upper and lower bounds for the identical transmission range r under the same network assumptions. They demonstrated that if rN is $\Omega(L \ln(L))$ ⁶, where N is the number of the nodes in the network of length L , then with probability one the network will be connected. In [47], the nodes, which are considered to be vehicles, perform a distributed algorithm to estimate the local density of the nodes in the network. The algorithm uses the mobility pattern of a node (vehicle) and is based on the stopping time of that node. The nodes adjust their transmission ranges according to the estimated density in a manner similar to that of [46]. Non-asymptotic results for a linear network of N nodes with density λ and with exponentially distributed distances between the nodes are also available [48], which indicate that the network will be connected with a probability greater than P_c , if $r \geq \frac{-\ln(1-P_c^{1/(N-1)})}{\lambda}$. Given the node density λ and the network length L , the value of r can thus be determined for a value of P_c arbitrarily close to one. In [49], an exact equation for the probability of having at most c clusters in a linear network with uniform node distribution and identical transmission range is given. When $c = 1$, this equation gives the probability of connectivity for the network. Papers [50, 51] studied the connectivity properties of one-dimensional wireless networks operating in Rayleigh fading and lognormal shadowing environments. The authors of [51] used the model in [50] to obtain connectivity properties for a VANET. They studied the effects of traffic flow and vehicle speed on the probability distribution of the connectivity distance. Both [50] and [51] considered identical transmission range for all the nodes of the network with homogenous Poisson distribution of the nodes.

More recently, in [1, 52–55], arbitrary node distribution in linear networks was considered. In [52], a stochastic traffic model for VANETs in signalized urban road systems is proposed. The proposed model composed of the fluid model and stochastic model. The fluid model is used to compute the average density of vehicles, while

⁵The transmission range assignment where all the nodes in the network transmit the data with the same transmission range.

⁶By definition, $f(x) = \Omega(g(x))$ if and only if $g(x) = \mathcal{O}(f(x))$.

the stochastic model takes into account the random behavior of individual vehicles. The authors validated the non-homogeneous Poisson distribution of the nodes on a road against empirical data in Central London. The same authors in [53] analyzed the connectivity dynamics of vehicular ad-hoc networks in the same network model. They determined the probability that the communication network is connected, i.e., each node can communicate with every other node through a multi-hop path, while all the nodes have identical transmission ranges. In [54], an arbitrary node distribution was considered, and it was shown that for a continuous density function $f(x)$, where the number of nodes in the network, N , goes to infinity, $\tau_{f,N} = \frac{1}{\inf\{f(x)\}} \frac{\log(N)}{N}$ is the critical transmission range, where $\inf\{\cdot\}$ denotes the infimum value. This means that if every node in the network has a transmission range greater than $\tau_{f,N}$, the network will be connected with probability one. It was further shown in [55] that for the networks with vanishing density functions a strong threshold does not exist. In [1], the authors considered a linear ad hoc network, where all the nodes used an identical transmission range. For an arbitrary spatial distribution of network nodes,⁷ they calculated the exact probability of connectivity of any location in the network to the source, which is assumed to be located at one end of the network.

2.3 Broadcasting Features Analysis

Aside from transmission range assignment, some works study the features of broadcasting mechanisms in wireless networks. In [26], the authors considered a linear wireless networks with uniform node distribution and identical transmission range assignment. For this network, they first obtained an approximation of the expected value and the standard deviation of the maximum possible distance covered in a single hop. Then they used these values to estimate the expected value and the standard deviation of the maximum coverage with multiple hops, which they modeled by a Gaussian random variable. Furthermore, they used this random variable to give an estimate of the hop-count distribution of an arbitrary location of the network.

To the best of our knowledge, there are a limited number of works studying hop-count probabilities in wireless networks [27, 56]. An estimation for the hop-count pmf in one-dimensional networks with identical transmission range assignment for

⁷The model used in [1] is implicitly based on Poisson distribution of network nodes with arbitrary density.

all the uniformly distributed nodes of the network is presented in [56]. In [27], the authors analyzed the performance of *Minimum Connected Dominating Set* (MCDS) broadcasting in one-dimensional networks, with identical transmission range for all the nodes, in which only the nodes in the MCDS⁸ retransmit the data. They derived the pdf of the maximum physical distance from the broadcast source reached by the k^{th} hop. They used the derived recursive formulas to obtain the mean distance and the number of nodes reached after a given number of hops, and the mean number of hops before the broadcast stops. Contrary to these papers, our analysis results presented here hold for more general node distributions and in cases where nodes do not use identical transmission ranges.

Our proposed localization protocol is range-free. Range-free algorithms do not need to measure the distance or angle information between unknown nodes and landmarks [11]. Range-based algorithms, on the other hand, must have the ability to measure the range of wireless signal transmissions. Some well-known ranging techniques are *received signal strength indication* (RSSI) and *time of arrival* (TOA) (or *time difference of arrival* (TDOA)) [57]. Although range-based protocols are generally more accurate than range-free protocols, their performance depends on the local distance measurement error of the ranging approach they use [58]. It is shown that for some cases, the local measurement error can be in the order of the transmission range of the nodes [59]. In [60], the authors calculated this error to be around 25% of the transmission range for the IEEE 802.15.4a UWB protocol analyzed in [61]. Even for transmission ranges as small as 0.6 to 18 meters, the local distance measurement error can be in the range of 0.5 to 10 meters [62]. In RSSI, the strength of the received signal is used to calculate its distance from the transmitter [63]. In TOA the distance between two nodes is calculated using the time the signal takes to propagate from one point to another. In TDOA, radio signal and ultrasound pulses are sent simultaneously, and using the time difference of the arrivals of the two signals, the distance between the sender and the receiver is calculated [57].

One of the most frequently cited methods that use the hop-count information for localization is DV-Hop [28] which is range-free. In this method, the anchors (aka

⁸The MCDS is the subset of connected nodes with minimum cardinality such that each node in the network is connected to a node in this subset. For a one-dimensional network, the MCDS can be recursively obtained by starting from the source and including in the MCDS, step by step, the farthest node within the transmission range of the previously inserted node.

reference nodes, beacon nodes, landmarks), which are either deployed at known locations or equipped with GPS devices, know the distance between themselves. They estimate the average length for one hop as the ratio of the distance to another anchor to the total number of hops needed to reach that anchor. The network nodes use these average hop distances to estimate their location (knowing the location of the anchors, and the hop-count to receive the message from each anchor).

Recently, in [30], a range-free approach, called DV-RND, based on DV-Hop is proposed. In this approach, instead of estimating the distance between two neighboring nodes, the authors defined the neighborhood distance between the two nodes. The neighborhood distance depends on the number of nodes covered by both nodes, and the number of nodes within each node's range. They then calculated the regulated neighborhood distance (RND) of all the neighboring nodes in the network, where RND of two neighboring nodes is the mean of the neighborhood distance of those nodes. The accuracy of the calculated RND depends on the assumption that the nodes are uniformly distributed between the two neighboring nodes. Localization of a node is performed using the sum of all the RNDs of any two nodes between all the anchors and that node. As a result, the amount of RND information needed to be exchanged among the nodes to obtain their location is quite large.

It is shown in [64] that most of the range-free localization protocols that do not attempt to assign a distance to a hop, but use connectivity as a constraint when jointly localizing all nodes, such as Multi-Dimensional Scaling (MDS) and Curvilinear Component Analysis (CCA), perform poorly in the networks with linear topology.

The range-based methods proposed in [65, 66] take into account the inaccurate range measurements of a single hop transmission. The inaccuracies are characterized by modeling the range measurements as a set of pdf's. In [65], the errors in distance estimations are modeled as normal random variables. When an unknown node receives a packet from an anchor node, its location pdf can be calculated. By receiving more packets from different anchors, the pdf of the location of the node is refined. More recently, in [66], the computational complexity of the probabilistic approach of [65] was reduced by using the Fast Fourier Transform (FFT).

Hop-count analysis can also be useful in securing wireless networks via location verification methods [13]. The probabilistic location verification (PLV) algorithm proposed in [13] uses the pmf of the hop-count from a location to a verifier, which is located at a fixed location. The verifiers are nodes equipped with PLV, and they determine

plausibility of the claimed location of a node, and a central authority gives an appropriate trust level to those claimed locations based on the calculated plausibility probabilities. The authors of [13] used the hop-count formulas they obtained in [56], and their analysis provides an estimation for the hop-count pmf, and is not exact. Analyzing the hop related properties can be used in the end-to-end delay analysis of wireless networks [7, 8]. In [7], the end-to-end delay of wireless line networks with known node locations is analyzed. The end-to-end delay analysis for two-dimensional wireless networks with known or random (uniformly distributed) node locations is presented in [8]. Furthermore, hop related properties can be useful in securing wireless networks via location verification methods [13, 14]. The probabilistic location verification (PLV) algorithm proposed in [13, 14] takes advantage of the probabilistic dependence of the hop-count for a specific location and the distance between the source and that location. The verifier nodes, equipped with PLV, determine the plausibility of the claimed location of a node, and based on the calculated plausibility probabilities, a central authority gives an appropriate trust level to those claimed locations.

Chapter 3

Energy-Efficient Broadcasting for Linear Networks

In this chapter, we investigate the transmission range assignment for N wireless nodes located on a line (a linear wireless network) for broadcasting data from one specific node to all the nodes in the network with minimum energy. Our goal is to find a solution that has low time complexity and yet performs close to optimal. We propose an algorithm for finding the optimal assignment (which results in the minimum energy consumption) with time complexity $\mathcal{O}(N^2)$. An approximation algorithm with time complexity $\mathcal{O}(N)$ is also proposed. It is shown that, for networks with uniformly distributed nodes, the linear-time approximate solution obtained by this algorithm on average performs practically identical to the optimal assignment. Both the optimal and the near-optimal algorithms require the full knowledge of the network topology and are thus centralized. We also propose a distributed algorithm of negligible time complexity, i.e., with time complexity $\mathcal{O}(1)$, which only requires the knowledge of the adjacent neighbors at each wireless node. Our simulations demonstrate that the distributed solution on average performs almost as good as the optimal one for networks with uniformly distributed nodes.

In our formulation, we pose no limit on the number of hops and assume that the consumed energy by each node is a function of the node's transmission range. The system model and the problem formulation is described in Section 3.1. In Section 3.2, we propose our algorithms for energy-efficient broadcasting. The first algorithm finds a near-optimal solution to the problem with time complexity $\mathcal{O}(N)$. The second algorithm finds an optimal solution and has time complexity $\mathcal{O}(N^2)$. Note that both algorithms are significantly less complex than the algorithm of [33] which has a time

complexity $\mathcal{O}(N^3)$. Similar to the algorithms of [6, 25, 33–35, 44], both proposed algorithms require the full knowledge of the network topology and are thus centralized. Finally, the last algorithm proposed in this work is a distributed one with negligible time complexity of $\mathcal{O}(1)$. To be implemented in each node of the network, this algorithm requires only the knowledge of the adjacent neighbor(s) of the node. This is much easier to attain compared to the amount of knowledge required for the centralized algorithms, i.e., the full knowledge of the network topology. Furthermore, this simplifies the implementation of the algorithm in mobile scenarios, where nodes only need to track their two closest neighbors. In Section 3.3, we present simulation results on the performance of the proposed algorithms. The contents of this chapter have been published as [67].

3.1 System Model

We consider a set of N nodes placed on a line having indices $\{1, 2, \dots, N\}$ from left to right, and a specific node with index s among them as the source. The source node s broadcasts data to all the other nodes in the network. This is to be performed in an energy-efficient multi-hop fashion. To solve the minimum-energy broadcasting problem, we need to assign a *transmission range* to each node so that the total consumed energy is minimized. A node is assumed to have symmetric coverage on both sides up to its transmission range and any other node located in the transmission range of this node can receive the transmitted data. In this work, similar to [6, 25, 33–35, 44–48, 54, 55], we do not consider the effects of interference caused by wireless communication among the nodes, and shadowing and fading, and also the overhead of obtaining information about the nodes' locations.¹

A *range assignment* R is a function $R : \{1, \dots, N\} \rightarrow \mathbb{R}^+$, where $R(i)$ is the assigned transmission range to node i . We denote the consumed energy of the range assignment R by $cost(R)$ and assume that it can be calculated, up to a constant multiplicative

¹Ignoring the interference can be justified by assuming that a scheduling scheme would ensure that simultaneous interfering transmissions will not occur. For ignoring fading/shadowing, one can assume that the transmissions occur in an environment with no obstacles, where the signals experience negligible fading/shadowing. If fading/shadowing is not ignored, then each link between different nodes of the network would experience a different loss and thus a different relationship between the distance and the consumed energy. This will change the system model compared to the one discussed in this work and is beyond the scope of this correspondence.

factor, using the following equation:

$$\text{cost}(R) = \sum_{k=1}^N R^\alpha(k), \quad (3.1)$$

where α is the *path-loss exponent* whose value is normally between 2 and 6 [3]. By using the *Minimum Energy Range Assignment* (denoted by R^*), every node in the network will receive the data transmitted by the source node with the minimum possible cost.

3.2 Energy-Efficient Range Assignments

In a linear network, each node has at most two immediate neighboring nodes, one on each side. We call these two nodes the adjacent neighbors of a node. In the sequel, we refer to the adjacent neighbor that is further away from the source as *the next adjacent neighbor* of a node.

Let us first consider a case where the source node is on one end of the network. There is a trivial optimal solution for this case [33]. In the optimal solution, the transmission range of each node, except the node which is at the other end of the network with respect to the source node, is equal to the distance to its next adjacent neighbor. For example if $s = 1$, the optimal solution will be as following:

$$\begin{aligned} R^*(i) &= d(i, i+1) \quad \text{for } i = 1, \dots, N-1 \\ R^*(N) &= 0, \end{aligned} \quad (3.2)$$

where $d(i, i+1)$ is the distance between nodes i and $i+1$.

This result is obtained by using the fact that for any given set $\{a_1, a_2, \dots, a_W\}$ of positive numbers, where W is an arbitrary integer, and for any $\alpha \geq 2$, we have:

$$\left(\sum_{k=1}^W a_k \right)^\alpha \geq \sum_{k=1}^W a_k^\alpha. \quad (3.3)$$

Hence, in the rest of the chapter, we assume that the source node is not at one end of the network ($s \neq 1, N$). Since nodes 1 and N do not have next adjacent neighbors to send data to, their transmission range in the optimal solution will always be zero. We divide the whole set of nodes excluding nodes 1 and N into two sets \mathbb{L} and \mathbb{R} ,

where:

$$\mathbb{L} = \{i : 1 < i \leq s\} \text{ and } \mathbb{R} = \{i' : s \leq i' < N\}. \quad (3.4)$$

We denote s by s_L when it is in \mathbb{L} , and by s_R when it is in \mathbb{R} .

The following lemma forms the basis of the proposed algorithms. Its proof is by contradiction and straight-forward.

Lemma 1. *In the minimum-energy range assignment, the transmission range of a node i is either zero or greater than or equal to the distance between i and its next adjacent neighbor.*

The minimum possible positive range of node i , $M(i)$, can be calculated as:

$$M(1) = M(N) = 0, \quad M(i) = \begin{cases} d(i, i - 1) & \text{for } i \in \mathbb{L}, \\ d(i, i + 1) & \text{for } i \in \mathbb{R}. \end{cases} \quad (3.5)$$

For $i = s$, there are two values of $M(i)$ corresponding to s_L and s_R , respectively.

3.2.1 Near-Optimal Range Assignment with Linear Time Complexity

For networks with known topology, we can save energy by preventing some nodes from transmission. These are the nodes with receivers located in the transmission range of other nodes. Fig. 3.1 shows an example of this situation. Node b receives the data from s and as b needs to transmit the data at a power level that can reach its next adjacent neighbor (node c), the data also reaches nodes a and d . Hence $R(a) = R(d) = R(c) = 0$.

The focus of our proposed near-optimal algorithm is on finding the nodes that can save energy by not transmitting, while the other nodes only transmit at a power level that is needed for their next adjacent neighbors to receive the data.

The pseudo-code for this algorithm is given as Algorithm 2.1: Near-Optimal Linear-Time Algorithm, that has the output R^{NO} as the range assignment. In this algorithm, we find the nodes with maximum other-side coverage on each side, and we decide to have one of them sending, while keeping the unnecessary nodes from transmitting.

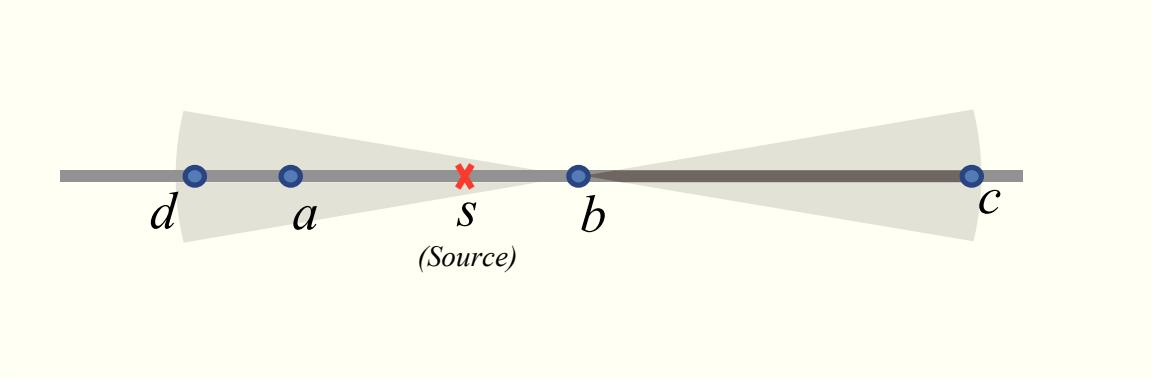


Figure 3.1: In an energy-efficient range assignment, some nodes do not need to transmit.

Algorithm 2.1 Near-Optimal Linear-Time Algorithm

Input: exact location of all the nodes

Output: R^{NO} (near-optimal transmission range assignment)

- 1: determine the sets \mathbb{L} and \mathbb{R}
- 2: **for** $i = 1$ **to** N **do**
- 3: $Cov(i) = M(i) - d(s, i)$
- 4: **end for**
- 5: $m_L = \arg \max_k \{Cov(k)\}$ for $k \in \mathbb{L}$
 $m_R = \arg \max_{k'} \{Cov(k')\}$ for $k' \in \mathbb{R}$
- 6: denote by l_L the left-most node in the opposite side coverage of m_R , and denote by l_R the right-most node in the opposite side coverage of m_L
- 7: calculate the costs:

$$cost_R = \begin{cases} cost^* & \text{if } l_R = s \\ \sum_{k=1}^{s_L} M^\alpha(k) + \sum_{k=l_R}^N M^\alpha(k) & \text{if } l_R > s \end{cases}$$

$$cost_L = \begin{cases} cost^* & \text{if } l_L = s \\ \sum_{k=1}^{l_L} M^\alpha(k) + \sum_{k=s_R}^N M^\alpha(k) & \text{if } l_L < s \end{cases}$$

where

$$cost^* = \max\{M(s_L), M(s_R)\}^\alpha + \sum_{\substack{k=1 \\ k \neq s}}^N M^\alpha(k)$$

8: **if** $cost_R \leq cost_L$ **then**

9:

$$R^{NO}(i) = \begin{cases} M(i) & i \in \{1, \dots, s_L, l_R, \dots, N\} \\ 0 & i \in \{s_R + 1, \dots, l_R - 1\} \end{cases}$$

10: **else**

11:

$$R^{NO}(i) = \begin{cases} M(i) & i \in \{1, \dots, l_L, s_R, \dots, N\} \\ 0 & i \in \{l_L + 1, \dots, s_L - 1\} \end{cases}$$

12: **end if**

In Steps 2 to 4 of Algorithm 2.1: Near-Optimal Linear-Time Algorithm, node s has two opposite side coverage values, each corresponding to one of its roles as s_L or s_R , i.e., $Cov(s_L) = M(s_L)$ and $Cov(s_R) = M(s_R)$. Note that in Step 6, l_R or l_L can be the source node s . In Step 7, the two costs correspond to two different ways of sending data from s to all the nodes in the network. The $cost_R$ ($cost_L$) is for the case where some nodes on the right (left) hand side of the source receive the data from m_L (m_R) and thus do not need to receive the data from their neighbor.

The following theorem is easy to prove.

Theorem 1. *The near-optimal algorithm has a time complexity of $\mathcal{O}(N)$.*

In this section, we assumed that the transmission range of a node is either equal to zero or equal to the distance to its next adjacent neighbor. For finding the optimal range assignment in the next section, we will use the fact that each node can have transmission range equal to zero or its distance to any node in the network (N possible values). Among all these possible assignments, we will prove that by searching a limited space the optimal solution can be found.

3.2.2 Optimal Range Assignment

The following theorem states some important facts about R^* .

Theorem 2. *The range assignment R^* satisfies the following conditions:*

1. *There exists at most one node (denoted by b_m) with transmission range greater than $M(b_m)$.*
2. *If b_m exists, it receives the data from s via the nodes in between s and itself (i.e., it does not receive the data via a node on the opposite side of the source).*
3. *If b_m exists, $R^*(b_m) \geq \max(d(b_m, l_R), d(b_m, l_L))$.*

Proof. We prove each statement as follows:

1. First we should note that the necessary condition for any node i to have $R^*(i) > M(i)$ is that by transmitting at this higher power, it must have a receiver on the opposite side of the source. Otherwise, using Equation (3.3), another range assignment with less consumed energy can be found. The proof then follows from the same arguments made in Lemmas 3 and 4 of [33]. This is done by substituting the concept of a *root-crossing* node in [33] with a node with $R^*(i) > M(i)$ in our context, and noticing that the *children* of a node in [33] are receivers of that node in our study. A node is called a root-crossing node in [33] if it has a receiver belonging to the other side of the source node (i.e., the side that it does not exist on). The children of a node in [33] denote all the receivers of that node.

Here we present a sketch of the proof. Assume there exist more than one node i with $R^*(i) > M(i)$. For those of the nodes with increased range (transmission range greater than their $M(.)$ value), which are on the same side of the source, we can assign $R^*(i) = M(i)$ except for one with the most number of receivers on the other side. By doing so, there will remain at most two nodes on opposite sides with increased range. Denote these nodes by a and b . At least one of them does not transmit beyond the farthest same-side receiver of the other one, as if they both can do, we will have $R^*(a) > R^*(b)$ and $R^*(a) < R^*(b)$. Consider node b as the one without any other-side receivers beyond the farthest same-side receiver of node a . In the process of transmitting data to node a , all the other-side receivers of node b will receive data. Therefore, increasing $R^*(b)$ from

$M(b)$ will not be beneficial. Hence, there exists at most one node (denoted by b_m) with transmission range greater than $M(b_m)$.

2. This is a direct result of Lemmas 3 and 4 of [33], which indicate that the optimal solution contains exactly one root-crossing node, i.e., if b_m exists, it is the only root-crossing node.
3. The proof for this part is by contradiction. Suppose that node b_m exists and is on the left side of the source (the proof for the right side is similar).

First, we prove that if $R^*(b_m) < d(b_m, l_R)$, the assignment R^* cannot be optimal. There can be more than one node having l_R in their transmission range. We denote the right-most node among those nodes on the left side of the source sending data to l_R by mR_L . If b_m is on the left side of mR_L , then since b_m receives the data from s via the nodes in between s and itself (including node mR_L), node l_R receives data from mR_L . The other possibility is for b_m to be on the right side of mR_L (in between nodes s and mR_L). In this case, since $M(mR_L) \geq d(mR_L, l_R) > d(b_m, l_R) > R^*(b_m)$, node b_m cannot transmit to the left receiver of node mR_L , so node mR_L still needs to transmit ($R^*(mR_L) = M(mR_L)$), which implies that l_R will be covered by the transmission from mR_L . Hence with $R^*(b_m) < d(b_m, l_R)$, node b_m cannot transmit data beyond node l_R on the right side of the source. Since all the nodes from s to l_R can receive the data from mR_L , node b_m has no receivers on its opposite side. Using Equation (3.3), we can easily show that by using the near-optimal range assignment R^{NO} , all the nodes can still receive the data with a lower cost, contradicting the optimality of the range assignment.

Now suppose that $R^*(b_m) < d(b_m, l_L)$. Note that since $d(b_m, l_R) \geq d(b_m, s)$, if node b_m is on the left side of node l_L , we will have: $R^*(b_m) < d(b_m, l_L) \leq d(b_m, s) \leq d(b_m, l_R)$ which results in a contradiction as previously discussed. We continue the proof for the case that node b_m is in between nodes s and l_L . Denote the left-most node on the right side of the source having l_L in its transmission range by mL_R . Also denote the last same-side receiver of node mL_R by ls_m (which is on the right side of the source). Since $d(b_m, ls_m) > M(mL_R) = d(mL_R, ls_m) \geq d(mL_R, l_L) > d(b_m, l_L) > R^*(b_m)$, node b_m cannot transmit to node ls_m , so node mL_R still needs to transmit ($R^*(mL_R) = M(mL_R)$). Node mL_R is a root-crossing node, and since there cannot be two root-crossing nodes

in the optimal solution, node mL_R does not receive the data via a node on the other side of the source (e.g., node b_m). Therefore node b_m has no receivers on the opposite side, and it has no receivers beyond node l_L , and since all the nodes in between nodes s and l_L receive the data from node mL_R , we can assign zero to the transmission range of node b_m . This results in reducing the energy consumption of R^* , which contradicts its optimality.

We showed that $R^*(b_m) \not\prec d(b_m, l_R)$ and $R^*(b_m) \not\prec d(b_m, l_L)$, therefore $R^*(b_m) \geq \max(d(b_m, l_R), d(b_m, l_L))$, and this completes the proof. \square

The following corollary is obtained based on the second and third propositions of Theorem 2.

Corollary 1. *If node b_m exists in R^* , then*

$$\text{cost}(R^*) = \begin{cases} (R^*(b_m))^\alpha + E_L + E_R & \text{if } b_m = s, \\ \sum_{k=b_m}^{s-1} d^\alpha(k, k+1) + \\ (R^*(b_m))^\alpha + E_L + E_R & \text{if } 1 < b_m < s, \\ \sum_{k=s}^{b_m-1} d^\alpha(k, k+1) + \\ (R^*(b_m))^\alpha + E_L + E_R & \text{if } s < b_m < N, \end{cases} \quad (3.6)$$

where $E_L = \sum_{j=1}^{r_L-1} d^\alpha(j, j+1)$ and $E_R = \sum_{j'=r_R}^{N-1} d^\alpha(j', j'+1)$ in which r_L and r_R denote the last left-side and right-side receivers of node b_m , respectively.

For reducing the time complexity of the optimal algorithm, we introduce two arrays and one matrix as follows. Let C_S be an array of size $N + 1$, where $C_S[i]$ ($i = 1, \dots, s_L, s_R, \dots, N$) is the cost of sending data from node s to node i via the nodes in between them. Construction of this array is performed by the following recursive

equations, having time complexity of $\mathcal{O}(N)$.

$$C_S[s_L] = C_S[s_R] = 0, \quad (3.7)$$

$$C_S[i] = \begin{cases} C_S[i+1] + M^\alpha(i+1) & 1 \leq i < s_L, \\ C_S[i-1] + M^\alpha(i-1) & s_R < i \leq N. \end{cases}$$

Similarly, we can compute another array C_E of size $N + 1$, where $C_E[i]$ ($i = 1, \dots, s_L, s_R, \dots, N$) is the cost of sending data from node i to the end node on its side, via the nodes in between them. Using the following recursive equations, we can construct C_E with time complexity $\mathcal{O}(N)$ in time.

$$C_E[1] = C_E[N] = 0, \quad (3.8)$$

$$C_E[i] = \begin{cases} C_E[i-1] + M^\alpha(i) & 1 < i \leq s_L, \\ C_E[i+1] + M^\alpha(i) & s_R \leq i < N. \end{cases}$$

Denote by LR a $2 \times N$ matrix in which column i contains the indices of the last same-side and other-side receivers of node i (denoted by rS_i and rO_i , respectively), when the transmission range of node i is equal to $\max(d(i, l_L), d(i, l_R))$. Note that constructing this matrix has a time complexity of $\mathcal{O}(N^2)$.

The pseudo-code for the optimal algorithm is given in Algorithm 2.2: Optimal Algorithm. The algorithm has the output R^* as the optimal range assignment. In this algorithm, using arrays C_S , C_E and the matrix LR , we find the node b_m (if it exists), and accordingly, we assign the ranges of the other nodes by just keeping the necessary ones to transmit. If no node b_m exists, this algorithm gives the same output as Algorithm 2.1: Near-Optimal Linear-Time Algorithm.

Algorithm 2.2 Optimal Algorithm

Input: exact location of all the nodes

Output: R^* (optimal transmission range assignment)

- 1: construct arrays C_S , C_E and the matrix LR
- 2: perform Algorithm 2.1: Linear-Optimal Linear-Time Algorithm
- 3: $COST = Cost(R^{NO})$
- 4: **for all** $b \in \{1, \dots, N\}$ **do**
- 5: read nodes rO_b and rS_b from matrix LR

```

6:    $cost(b, rO_b, rS_b) = C_S[b] + [\max(d(b, rO_b), d(b, rS_b))]^\alpha + C_E[rS_b] + C_E[rO_b]$ 
7:   if  $cost(b, rO_b, rS_b) < COST$  then
8:      $COST = cost(b, rO_b, rS_b)$ 
9:      $b_m = b$ ,  $rO = rO_b$  and  $rS = rS_b$ 
10:    end if
11:    denote the next adjacent neighbors of nodes  $rO_b$  and  $rS_b$  by  $nan(rO_b)$  and
         $nan(rS_b)$ , respectively
12:    if either  $nan(rO_b)$  or  $nan(rS_b)$  exists then
13:      select the one which is closer to node  $b$  (e.g.  $nan(rS_b)$ )
14:      replace the element of  $LR$  containing  $rS_b$  with  $nan(rS_b)$ 
15:      if  $d(b, nan(rO_b)) = d(b, nan(rS_b))$  then
16:        replace the element containing  $rO_b$  in  $LR$  with  $nan(rO_b)$ 
17:        Goto 5
18:    end if
19:  end if
20: end for
21: if no node  $b_m$  has been found then
22:    $R^* = R^{NO}$ 
23: else
24:    $R^*(b_m) = \max(d(b_m, rO), d(b_m, rS))$ 
25:    $R^*(i) = M(i)$  for all  $i$  from  $s$  up to  $b_m$ , and from  $rO$  and  $rS$  to both ends of
        the network correspondingly
26:    $R^*(k) = 0$  for all the remaining nodes  $k$ 
27: end if
 $cost(R^*) = COST$ 

```

Theorem 3. *Obtaining R^* by Algorithm 2.2: Optimal Algorithm, has time complexity $\mathcal{O}(N^2)$.*

3.2.3 Distributed Range Assignment

In the last algorithm, every node just knows the distances to its adjacent neighbors. The amount of required information in this case is much less than that of the previous scenarios. Since each node only requires local information, this algorithm can be implemented in a distributed manner.

In the proposed distributed algorithm, each node should wait till it receives data from one of its adjacent neighbors. It then transmits the data to its other neighbor. The range assignment R^D for this algorithm is thus:

$$\begin{aligned} R^D(s) &= \max\{M(s_R), M(s_L)\}, \\ R^D(i) &= M(i) \text{ for } i \neq s. \end{aligned} \quad (3.9)$$

If we assume that the N nodes are located on a line according to a uniform distribution, and independent of one another, then the distances between the adjacent nodes are independent and identically distributed (i.i.d.) random variables $\{D_1, \dots, D_{N-1}\}$, where D_i denotes the distance between nodes i and $i + 1$. If the network length and the number of the nodes go to infinity, these random variables have an exponential distribution, denoted by $\exp(\lambda)$ with the probability density function given by $\lambda e^{-\lambda x}$ for $x \geq 0$, where λ is the density of the nodes on the line. For a finite-length network, with a finite number of nodes, we approximate the inter-node distances to have the same exponential distribution. Under such assumptions, the expected total cost of the distributed range assignment can be approximated as follows.

$$\text{Expected Cost} \simeq \sum_{\substack{k=1 \\ k \neq s-1, s}}^{N-1} E[(D_k)^\alpha] + E[(\max\{D_{s-1}, D_s\})^\alpha] \simeq \frac{\alpha!}{\lambda^\alpha} \left(N - 1 - \frac{1}{2^\alpha} \right). \quad (3.10)$$

3.2.4 Identical Transmission Range

When the distances between the nodes are drawn i.i.d. from $\exp(\lambda)$, a simple solution to the transmission range assignment, for maintaining the network connectivity with a given probability P_c , is to assign an identical transmission range $R(P_c)$ to all the nodes in the network. This assignment must satisfy the following inequality:

$$R(P_c) \geq \frac{-\ln(1 - P_c^{1/(\lambda L - 1)})}{\lambda}. \quad (3.11)$$

Note that Equation (3.11) is similar to the one given in [48] where L is approximated by N/λ . By ignoring the term -1 in $\lambda L - 1$, and using the first two terms of the Taylor series of $P_c^{1/(L\lambda)}$ for variable $1/L$ in the neighborhood of zero, we obtain the

following approximation for the lower bound of the identical transmission range:

$$R_l(P_c) \approx \frac{\ln\left(\frac{-\lambda L}{\ln(P_c)}\right)}{\lambda}. \quad (3.12)$$

The range assignment in Equation (3.12) can be used when the only available information about the network is its density and length. Moreover, using this range assignment, there is no guarantee that the network will in fact be connected. On the other hand, the condition of the network being connected is in general stronger than the condition required for a specific source node in the network to broadcast its message to all the other nodes in the network.

3.3 Numerical Results

The simulation results are presented for networks with N nodes distributed uniformly and independently over a line of length L . In simulations, we assume $\lambda = N/L$ and $\alpha = 2$. For each simulation point corresponding to a given density λ , 10,000 random networks are generated. For the results comparing the different algorithms, we run the algorithms on exactly the same networks and obtain the average of the total consumed energy over the 10,000 networks.

In Fig. 3.2, we compare the energy consumption of the identical transmission range assignment of Section 3.2.4 (for different P_c values) with those of the near-optimal range assignment of Section 3.2.1, the optimal range assignment of Section 3.2.2 and the distributed range assignment given in Section 3.2.3 for a network with $L = 5000$ meters. As can be seen in Fig. 3.2, in general, the total consumed energy decreases as the density of the nodes increases. This can be explained by Equation (3.3), where nodes with closer distance to each other can communicate with less energy over a given distance compared to nodes that are further apart. The proposed assignments significantly outperform the identical transmission range assignment, even for a P_c value as small as 0.85. Interestingly, both the linear-time near-optimal algorithm and the distributed algorithm perform practically the same as the optimal algorithm over the whole range of network densities. For the distributed algorithm, the simulation results and the analytical results from Equation (3.10) are almost identical.

To obtain a more detailed picture of the relative energy consumption of the proposed algorithms, we have plotted the histogram of the normalized difference between their

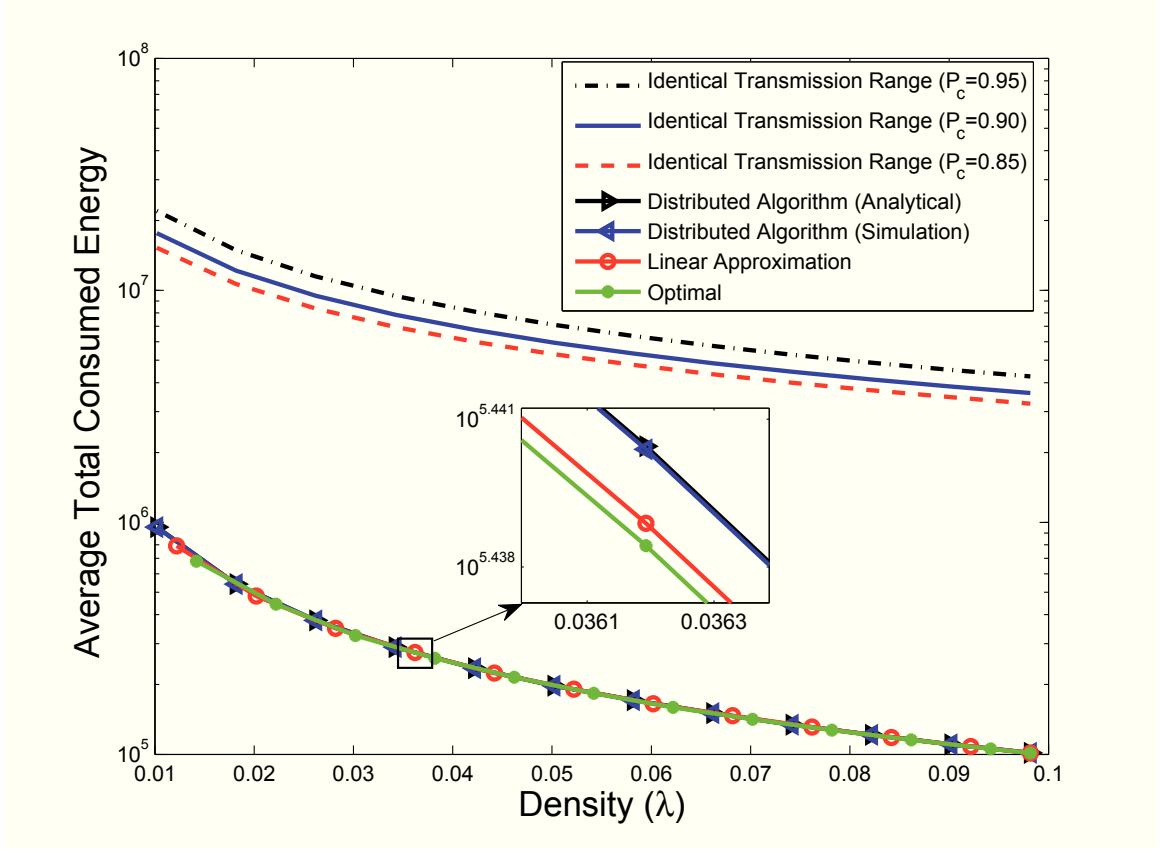


Figure 3.2: Comparison of the energy consumption of different range assignments.

consumed energy for $\lambda = 0.03$ in Fig. 3.3. The normalized difference between the energy consumption of range assignments R_1 and R_2 is defined as:

$$\frac{\max(\text{cost}(R_1), \text{cost}(R_2)) - \min(\text{cost}(R_1), \text{cost}(R_2))}{\min(\text{cost}(R_1), \text{cost}(R_2))} . \quad (3.13)$$

Fig. 3.3 demonstrates that the normalized difference between the energy consumption of the three proposed algorithms is rather small (less than 10%) for the simulated cases. The simple distributed algorithm and the linear near-optimal algorithm consume at most 9% and 6% more energy than the complex optimal algorithm, respectively.

Although the difference between the energy consumption of the more complex range assignment algorithms compared to less complex ones is rather small in networks with uniformly distributed nodes, there exist network topologies where such differences

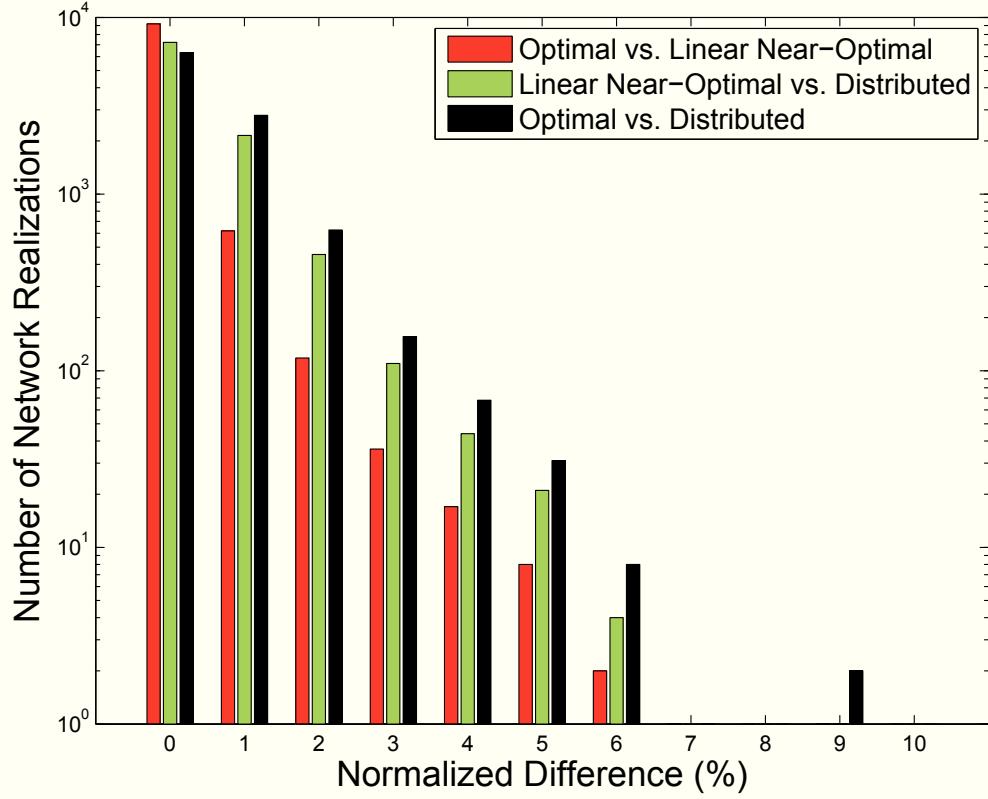


Figure 3.3: The histogram of the normalized differences between the energy consumptions of the proposed optimal, linear near-optimal and distributed algorithms for $\lambda = 0.03$.

are large. Fig. 3.4 illustrates two examples of such topologies. In both examples, we assume that $\epsilon_1, \epsilon_2 \ll r_1, r_2$, and that $\epsilon_1 \simeq \epsilon_2$, for simplicity. In the network given in Fig. 3.4(a), assuming that $r_1 \geq r_2 + \epsilon_1 + \epsilon_2$, and $r_1 \simeq r_2$, the normalized difference in energy consumptions of the linear near-optimal and distributed algorithms, calculated by (3.13), is about 100%.

In the network given in Fig. 3.4(b), assuming that $r_1 \leq r_2 + \epsilon_1 + \epsilon_2$ and $r_1 + \epsilon_1 \geq r_2 + \epsilon_2$, in the optimal range assignment, node s sends data to node c , that in turn transmits data to nodes a , b and d . In the linear near-optimal range assignment, node s transmits data to node d , and thus also covers nodes c and b . Node b then transmits data to node a . Assuming $r_1 \simeq r_2$, the normalized difference between the consumed energy of the optimal and the linear near-optimal algorithms is about 100%.

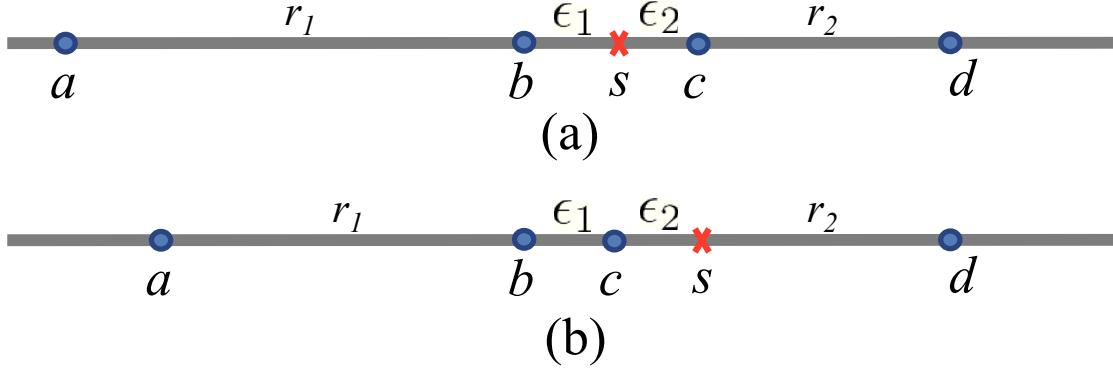


Figure 3.4: Network topologies where the difference between the consumed energy of different algorithms (linear near-optimal vs. distributed in (a), and optimal vs. linear suboptimal in (b)) can be large.

Finally in Fig. 3.5, we show the histogram of the distance $d(b_m, s)$ between node b_m in the optimal solution and the source node for 10,000 generated networks with $\lambda = 0.03$. This corresponds to the same scenarios used for Figs. 3.2 and 3.3. The histogram of Fig. 3.5 shows that for the majority of cases, node b_m does not exist. In most of the remaining cases, the distance $d(b_m, s)$ is relatively small, with the maximum distance less than 30% of the length of the network. These are the main reasons behind the small difference among the energy consumption of the three proposed range assignments in the simulated scenarios as reflected in Figs. 3.2 and 3.3. A careful inspection of Figs. 3.3 and 3.5 also demonstrates that although there exist some cases where $d(b_m, s)$ is relatively large (about 0.3L), the difference in energy consumption is relatively low (less than 10%). The reason is that in such cases although energy is saved in the optimal solution through the coverage of opposite side nodes by node b_m , node b_m itself consumes a large amount of energy due to its large transmission range.

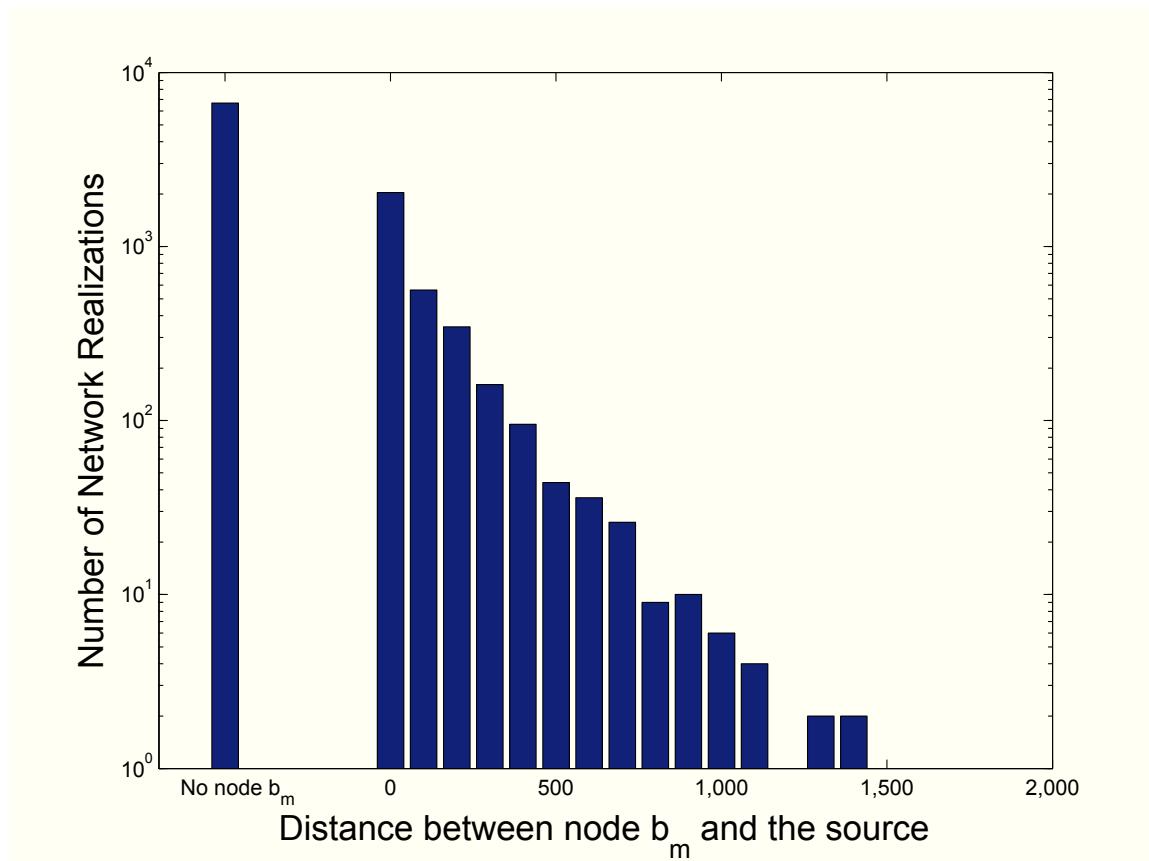


Figure 3.5: The histogram of the distance between node b_m and the source node for $\lambda = 0.03$.

Chapter 4

Energy-Efficient Broadcasting for Cross Networks

In Chapter 3, we solved the minimum-energy broadcasting problem for linear networks. In this chapter, we propose solutions for the same problem, but for cross networks, where the nodes are located on two perpendicular lines. Our solutions consist of an algorithm which finds the optimal range assignment in polynomial time ($\mathcal{O}(N^{12})$), a near-optimal algorithm with linear time complexity ($\mathcal{O}(N)$), and a distributed algorithm with time complexity $\mathcal{O}(1)$. To the best of our knowledge, this is the first study presenting an optimal solution for the minimum-energy broadcasting problem for a 2-D network (with cross configuration). We compare our algorithms with the broadcast incremental power (BIP) algorithm, one of the most commonly used methods for solving this problem with time complexity $\mathcal{O}(N^2)$. We demonstrate that our near-optimal algorithm outperforms BIP, and that the distributed algorithm performs close to it. Moreover, the proposed distributed algorithm can be used for more general two-dimensional networks, where the nodes are located on a grid consisting of perpendicular line-segments. The performance of the proposed near-optimal and distributed algorithms tend to be closer to the optimal solution for larger networks. We prove that the optimal, near-optimal and distributed algorithms perform the same in the asymptotic regime where N tends to infinity.

4.1 System Model

We consider a set of N nodes (denoted by \mathcal{N}) located on a cross, including a specific node, s , as the source. We assume that the exact location of the nodes is known. The

source node can be located anywhere on the cross and broadcasts the data to all the other nodes in the network. This is to be performed in an energy-efficient multi-hop fashion.

We assume that all the nodes are equipped with an omnidirectional transmitter. Moreover, in our study, similar to [6, 24, 25, 31, 36, 40–43, 67], we do not consider other issues such as channel contention and interference.

To solve this broadcasting problem, we need to assign a *transmission range* to each node so that the total consumed energy is minimized, while the data is delivered to all the nodes of the network. A *transmission range assignment* R is a function $R : \mathcal{N} \rightarrow \mathbb{R}^+$, where $R(i)$ is the assigned transmission range to node $i \in \mathcal{N}$. We denote the consumed energy of the range assignment R by $\text{cost}(R)$ and assume that it can be calculated, up to a constant multiplicative factor, using the following equation:

$$\text{cost}(R) = \sum_{k \in \mathcal{N}} R^\alpha(k), \quad (4.1)$$

where α is the *path-loss exponent* whose value is normally between 2 and 6 [3]. By using the *Minimum-Energy Range Assignment* (a.k.a, Optimal Range Assignment), denoted by R^* , every node in the network will receive the data transmitted by the source node with the minimum possible cost.

A node is assumed to have circular coverage up to its transmission range, and any other node located in the transmission range of this node can receive the transmitted data. Some definitions and notations are presented in Fig. 4.1.

According to the circular transmission range of the nodes, as shown in Fig. 1, for every node $a \in \mathcal{N}$, we have $R(a) = \text{cov}_{\text{same}}^R(a) \geq \text{cov}_\perp^R(a) \geq \text{cov}_{\text{oppo}}^R(a)$.

In the following, we provide some definitions needed throughout the chapter.

Definition 1. A cross network has five **segments**, as shown in Fig. 4.2. Segments I, III, IV and V are half-lines, while Segment II is the line-segment bounded by the source node and the intersection of the two perpendicular lines. For any node $a \in \mathcal{N} \setminus \{s\}$,¹ we denote the segment on which it is located by \mathcal{S}_a .²

¹For any two sets \mathcal{A} and \mathcal{B} , notation $\mathcal{A} \setminus \mathcal{B}$ denotes the set of all elements which are members of \mathcal{A} but not members of \mathcal{B} .

²By definition, the source node is not on any segment.

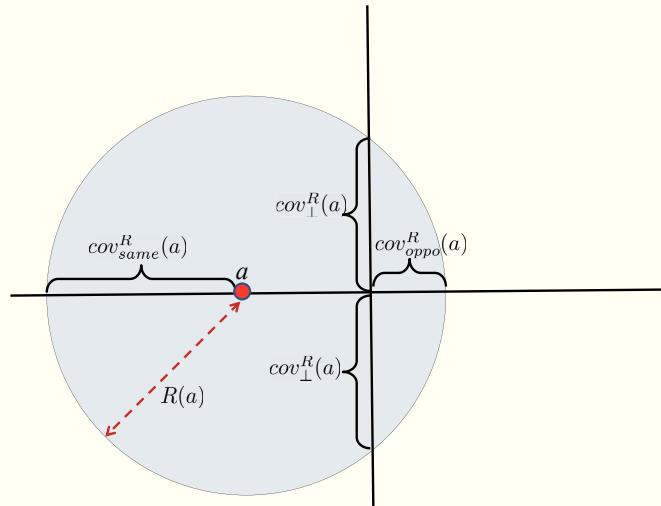


Figure 4.1: Some definitions/notations for networks on a cross.

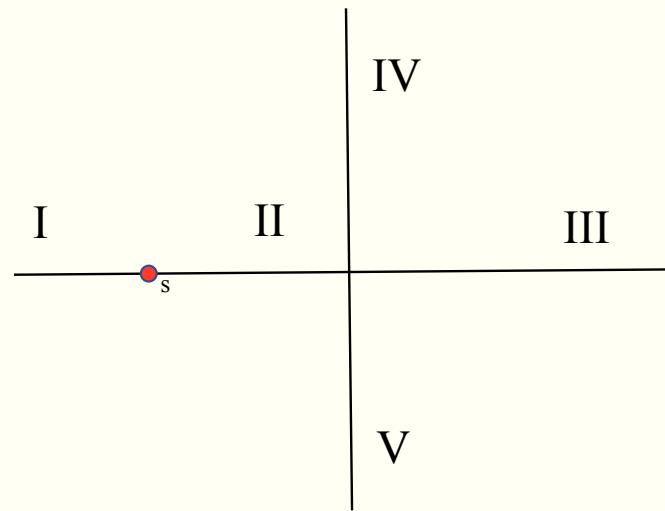


Figure 4.2: Segmentation of cross networks.

Definition 2. On each segment, the closest node to the source node is called the first node of that segment. The last node of a segment is the node farthest away from the source on that segment. We denote the first node and the last node of any segment \mathcal{S} by $f_{\mathcal{S}}$ and $l_{\mathcal{S}}$, respectively.³

³If Segment II is empty, we assume that node s takes all the functionalities of node l_{II} .

We denote the set of nodes $\mathcal{N} \setminus \{s, l_{\text{II}}, f_{\text{III}}, f_{\text{IV}}, f_{\text{V}}\}$ by $\hat{\mathcal{N}}$.

Definition 3. *Node a is after node b on the same segment, if $\mathcal{S}_a = \mathcal{S}_b$ and $d(s, a) > d(s, b)$. For these two nodes, we say node b is before node a .*

Definition 4. *For any node $a \in \mathcal{N} \setminus \{s\}$, we call the first node after a on \mathcal{S}_a , **the next adjacent neighbor** of node a , and we denote it by n_a . Furthermore, we define $M(a) = d(a, n_a)$. If node a is the last node on its segment, as it does not have a next adjacent neighbor, we define $M(a) = 0$.*

Fig. 4.3 illustrates the next adjacent neighbor of some nodes in a cross network.

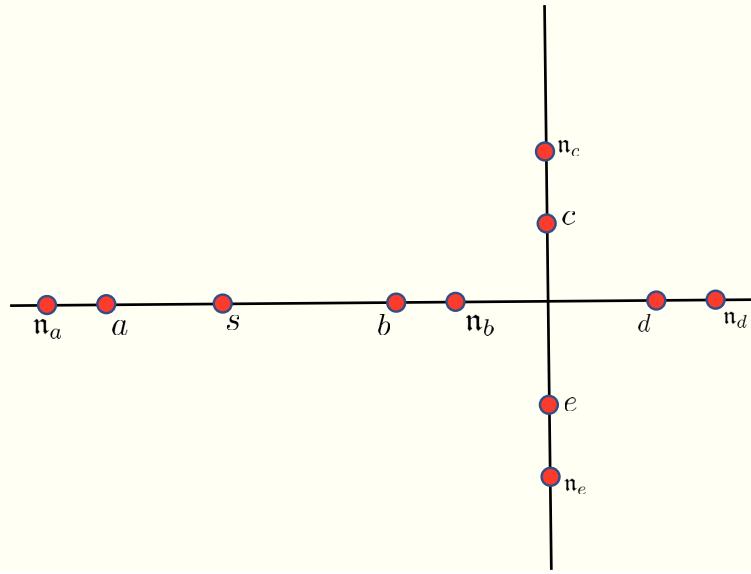


Figure 4.3: The next adjacent neighbors of nodes a, b, c, d and e .

Definition 5. *For assignment R , we say that node $a \in \mathcal{N} \setminus \{s\}$ has **increased (transmission) range**, if $R(a) > M(a)$.*

Definition 6. *In a cross network utilizing transmission range assignment R , for any node $a \in \mathcal{N} \setminus \{s\}$, we call the set of all nodes located after node a on \mathcal{S}_a , that are within the transmission range of node a , **the same-segment receivers** of node a . We call the set of nodes on segments other than \mathcal{S}_a that are within the transmission range of node a , **the other-segment receivers** of node a . The union of these two sets for node a is called **the receivers** of node a .*

Definition 7. In a cross network utilizing transmission range assignment R , for any two nodes $a, b \in \mathcal{N}$, node b receives the data via a path (starting from the source node) containing node a , if for node b to receive the data, node a has to transmit with $R(a)(\neq 0)$. In other words, if node a does not transmit, node b will not receive the data (at all). We show this relationship by $b \xleftarrow{R} a$. We show the case where node b can receive the data even if node a does not transmit by $b \not\xleftarrow{R} a$.

This concept is important to note, because a node has to receive the data first to be able to transmit it to other nodes. So, if $b \xleftarrow{R} a$, first node a has to transmit the data with $R(a)$, then node b will be able to transmit it to other nodes.

Also note that for nodes a and b , the relation $b \xleftarrow{R} a$ does not necessarily mean that node b receives the data in just one hop from node a . It means that the data travels through node a to get to node b , and this is the only way for node b to receive it.

Another important property of this concept is that for any two nodes $a, b \in \mathcal{N}$, we can have either $a \xleftarrow{R} b$ or $b \xleftarrow{R} a$, and not both. But we can have $a \not\xleftarrow{R} b$ or $b \not\xleftarrow{R} a$ or both.

Definition 8. In a cross network utilizing transmission range assignment R , we call those receivers b of node $a \in \mathcal{N}$, **the intended receivers** of a , if for all of them, we have $b \xleftarrow{R} a$. We denote the set of intended receivers of node a with transmission range $R(a)$ by \mathfrak{I}_a^R .

4.2 Proposed Range Assignments

In the following, we explain our proposed range assignments, starting from the optimal one, followed by the near-optimal assignment and finally the distributed one.

4.2.1 Optimal Range Assignment

One of our main contributions is to prove that in the optimal range assignment, there exists a small (and independent from N) set of nodes with increased transmission range, and the other nodes have either 0 or the distance to their next adjacent neighbor as their transmission range. Also, we provide an algorithm with polynomial time complexity in N to find the optimal assignment for the nodes on a cross.

An upper bound of eight on the number of nodes with increased transmission range in the optimal range assignment is established in the following theorem.

Theorem 4. *In the optimal range assignment for a cross network, denoted by R^* , there exist at most three nodes with increased transmission range in $\hat{\mathcal{N}}$. All the other nodes in $\hat{\mathcal{N}}$ have either 0 or the distance to their next adjacent neighbor as their transmission range. The optimal transmission range of a node in the set $\{s, l_{II}, f_{III}, f_{IV}, f_V\}$ can be equal to its distance to any other node in the network.*

For proving this theorem, we first introduce some lemmas.

Lemma 2. *For any node $a \in \mathcal{N}$, if $R^*(a) \neq 0$, then $\mathfrak{I}_a^{R^*} \neq \emptyset$.*

Lemma 3. *For any given set $\{a_1, a_2, \dots, a_W\}$ of positive numbers, where W is an arbitrary integer, and for any $\alpha \geq 2$, we have:*

$$\left(\sum_{k=1}^W a_k \right)^\alpha \geq \sum_{k=1}^W a_k^\alpha. \quad (4.2)$$

Corollary 2. *To transmit data on a line from node a to node b with minimum energy consumption, the data has to be transmitted hop by hop using the nodes in between transmitter and receiver. This means that node a transmits the data to node n_a , and node n_a transmits to its next adjacent neighbor, and so forth till the data reaches node b . We denote this transmission scheme by $a \dashrightarrow b$, where all the nodes from a up to b use their M values as their transmission ranges.*

Lemma 4. *In a cross network utilizing transmission range assignment R , if node $a \in \mathcal{N} \setminus \{s\}$ receives the data, all the nodes before it on its segment have also received the data.*

Proof. This is a direct result of the circular transmission range assumption of the nodes. \square

Corollary 3. *In a cross network utilizing transmission range assignment R , for nodes $a, b \in \mathcal{N} \setminus \{s\}$ where $\mathcal{S}_a = \mathcal{S}_b$, and node b is after node a , we have $a \xleftarrow{R} b$.*

Lemma 5. *In a cross network utilizing transmission range assignment R , for nodes $a \in \mathcal{N} \setminus \{s\}$ and $c \in \mathcal{N}$, if $a \xleftarrow{R} c$, then we have $b \xleftarrow{R} c$ for all nodes b after node a on \mathcal{S}_a .*

Proof. The proof is by contradiction. Suppose for two nodes a and b on the same segment, where node b is after node a we have $a \xleftarrow{R} c$, but $b \not\xleftarrow{R} c$. This means that if node c does not transmit the data, then node a will not receive it, but node b receives it from another path. This contradicts Lemma 4. \square

The contraposition of Lemma 5 gives us the following corollary.

Corollary 4. *In a cross network utilizing transmission range assignment R , if $b \not\leftrightarrow^R c$ for node $b \in \mathcal{N} \setminus \{s\}$, then we have $a \not\leftrightarrow^R c$ for all nodes a before node b on its segment.*

Throughout the chapter, we implicitly use the following lemma.

Lemma 6. *In a cross network utilizing transmission range assignment R , any node $a \in \mathcal{N}$ is the intended receiver of at most one other node.*

Lemma 7. *In a cross network utilizing transmission range assignment R , if for nodes $a, b \in \mathcal{N}$ we have $b \not\leftrightarrow^R a$, then the intended receivers of node a on any segment \mathcal{S} (if any exists) are after the last receiver of b on \mathcal{S} , which is denoted by $r_{b,\mathcal{S}}^R$ (if this node exists).*

Proof. Node b receives the data via a path that does not contain node a , so if node a does not transmit, node b and (according to Lemma 4) all the nodes from $f_{\mathcal{S}_b}$ up to b on segment \mathcal{S}_b receive the data. Having the circular transmission range of the nodes in mind, we can see that when node b transmits the data, all the nodes from $f_{\mathcal{S}}$ (if $\mathcal{S} \neq \mathcal{S}_b$) or node b (if $\mathcal{S} = \mathcal{S}_b$) up to $r_{b,\mathcal{S}}^R$ on segment \mathcal{S} receive the data. So all of the nodes from $f_{\mathcal{S}}$ to $r_{b,\mathcal{S}}^R$ on segment \mathcal{S} , receive the data even if node a does not transmit. Hence they are not in the set \mathfrak{I}_a^R . Therefore, for node a to have some intended receivers on segment \mathcal{S} , they have to be after node $r_{b,\mathcal{S}}^R$. \square

Lemma 8. *For every node $a \in \hat{\mathcal{N}}$, $R^*(a)$ is either zero or greater than or equal to $M(a)$. Furthermore, if this node has an increased transmission range, i.e., if $R^*(a) > M(a)$, it must have at least one intended receiver on a segment other than \mathcal{S}_a .*

Proof. According to Lemma 2, node a does not have $0 < R^*(a) < M(a)$, if it has no intended receivers. Using contradiction, suppose $R^*(a) < M(a)$, while it has some intended receivers. This implies that node a does not have any receivers on \mathcal{S}_a , so all the nodes in $\mathfrak{I}_a^{R^*}$ are on other segments. We denote the farthest receiver of node a when it transmits with $R^*(a)$ by $r_a^{R^*}$. For the node k defined below, we have

$$d(a, r_a^{R^*}) > d(k, r_a^{R^*}).$$

$$k = \begin{cases} s, & \text{if } a \text{ is on Segment I;} \\ s, & \text{if } a \text{ is on Segment II;}^4 \\ f_{\text{III}}, & \text{if } a \text{ is on Segment III;} \\ f_{\text{IV}}, & \text{if } a \text{ is on Segment IV;} \\ f_{\text{V}}, & \text{if } a \text{ is on Segment V.} \end{cases} \quad (4.3)$$

Since node a does not have any same-segment receivers, if we have another assignment R with $R(a) = 0$, $R(k) = \max\{R^*(k), d(k, r_a^{R^*})\}$, and $R(i) = R^*(i)$, $\forall i \in \mathcal{N} \setminus \{k, a\}$, all the nodes will receive the data, but less energy will be consumed. This contradicts the optimality of R^* .

The proof of the next part is also by contradiction. Suppose $R^*(a) > M(a)$, but all of the intended receivers of node a are on its own segment. According to Corollary 2, and since all the intended receivers of node a are on the same line, by using $a \dashrightarrow r_{a, S_a}^{R^*}$ the data will be delivered to all the intended receivers of node a with less energy. In this case node a does not have increased range, which contradicts our assumption. \square

Lemma 9. *In a cross network utilizing transmission range assignment R , a node on one of the Segments III, IV or V does not have intended receivers on Segment II.*

Proof. As a result of the circular transmission range assumption of the nodes, if a node on one of the Segments III, IV or V receives the data, all the nodes on Segment II have already received the data. \square

Lemma 10. *Consider a cross network that utilizes a range assignment R , for which $R(a) \neq 0$ and $R(b) \neq 0$ for nodes $a, b \in \hat{\mathcal{N}}$ (node a being closer to source, if they are on the same segment). If node a has an intended receiver after the last same-segment receiver of node b (node r_{b, S_b}^R), then assignment R is not the optimal assignment.*

Proof. The proof is by contradiction. Suppose $R = R^*$, and node $a \in \hat{\mathcal{N}}$ has an intended receiver after the last same-segment receiver of node $b \in \hat{\mathcal{N}}$. All the same-segment receivers of node b are within the transmission range of node a . Three cases

⁴If node a is located on Segment II, it must have $R^*(a) > M(a)$ to cover some nodes on Segments III, IV or V. Hence, if $0 < R^*(a) < M(a)$, all the other-segment receivers of a are located on Segment I.

may exist:

- 1) Nodes a and b are on the same segment and $cov_{same}^R(a) \geq cov_{same}^R(b)$.
- 2) Node a is on a segment aligned to the segment of node b .
- 3) Node a is on a segment perpendicular to the segment of node b .

Case (1): According to Corollary 3, $a \xrightarrow{R} b$. The whole transmission circle of node b is within the range of node a . So all the receivers of b can receive the data from a , i.e., $\mathcal{J}_b^R = \emptyset$, which contradicts the optimality of assignment R (Lemma 2).

Case (2): For node a to have an intended receiver on \mathcal{S}_b , we can not have $\mathcal{S}_a = \text{III}$ and $\mathcal{S}_b = \text{II}$ (according to Lemma 9). For the other cases, node a having an intended receiver after r_{b,\mathcal{S}_b}^R , results in having the whole circular range of b within the range of a . Hence, all the receivers of b can receive the data from a . To be sure about the reception of data by node a , we define node k as follows:

$$k = \begin{cases} s, & \text{if } b \text{ is on Segment I, and } a \text{ is on either Segments II or III;} \\ f_{\text{III}}, & \text{if } b \text{ is on Segment III, and } a \text{ is on Segment II;} \\ s, & \text{if } b \text{ is on Segment III, and } a \text{ is on Segment I;} \\ f_{\text{IV}}, & \text{if } b \text{ is on Segment IV, and } a \text{ is on Segment V;} \\ f_{\text{V}}, & \text{if } b \text{ is on Segment V, and } a \text{ is on Segment IV.} \end{cases} \quad (4.4)$$

We can have another assignment, R' , with $R'(b) = 0$, $R'(i) = R(i)$, $\forall i \in \mathcal{N} \setminus \{k, b\}$, and $R'(k) = \max\{d(k, ro_b^R), R(k)\}$, where ro_b^R denotes the farthest other-segment receiver of node b when it transmits with $R(b)$. Note that since $k \xrightarrow{R} b$ (Corollary 3), node b has no effect on the delivery of data to node k . Using R' all the nodes will receive the data with less energy, which contradicts the optimality of assignment R . Case (3): Similar to the previous case, and according to Lemma 9, for node a to have an intended receiver on \mathcal{S}_b , we can not have $\mathcal{S}_a = \text{IV}$ or $\mathcal{S}_a = \text{V}$, while $\mathcal{S}_b = \text{II}$. Define

node k' as follows:

$$k' = \begin{cases} s, & \text{if } b \text{ is on Segment I, and } a \text{ is on either Segments IV or V;} \\ f_{\text{III}}, & \text{if } b \text{ is on Segment III, and } a \text{ is on either Segments IV or V;} \\ f_{\text{IV}}, & \text{if } b \text{ is on Segment IV, and } a \text{ is on one of the Segments I, II or III;} \\ f_{\text{V}}, & \text{if } b \text{ is on Segment V, and } a \text{ is on one of the Segments I, II or III.} \end{cases} \quad (4.5)$$

We can have another assignment, R' , with $R'(b) = 0$, $R'(i) = R(i)$, $\forall i \in \mathcal{N} \setminus \{k', b\}$, and $R'(k') = \max\{d(k', ro_b^R), R(k')\}$, where ro_b^R denotes the farthest other-segment receiver of node b when it transmits with $R(b)$. Note that since $k' \not\leftrightarrow b$ (Corollary 3), node b has no effect on the delivery of data to node k' . Using R' all the nodes will receive the data with less energy, which contradicts the optimality of assignment R . \square

The following corollary is a direct result of Lemma 10.

Corollary 5. *If node $a \in \hat{\mathcal{N}}$ does not have increased transmission range, and node n_a receives the data from another node, then $R^*(a) = 0$.*

Lemma 11. *Consider a cross network that utilizes the optimal assignment R^* . For nodes $a, b \in \hat{\mathcal{N}}$, if $R^*(a) > M(a)$ and $R^*(b) > M(b)$, we have either $a \xleftarrow{R^*} b$ or $b \xleftarrow{R^*} a$.*

Proof. The proof is by contradiction. Suppose nodes a and b have increased range, and we have $a \not\leftrightarrow b$ and $b \not\leftrightarrow a$. Since nodes a and b have increased range, according to Lemma 8, they must have intended receivers on segments other than their own. According to Lemmas 7 and 10, none of them has other-segment intended receivers on \mathcal{S}_a and \mathcal{S}_b . For any of the other segments, e.g., \mathcal{S} , using Lemma 7, the intended receivers of node a on \mathcal{S} have to be after $r_{b,\mathcal{S}}^{R^*}$ and the intended receivers of node b on \mathcal{S} have to be after $r_{a,\mathcal{S}}^{R^*}$, which is impossible and contradicts our assumption. \square

Lemma 12. *In a cross network, utilizing R^* , a node in $\hat{\mathcal{N}}$ with increased transmission range on Segment II must have at least one intended receiver on Segment I.*

Proof. The proof is by contradiction. Suppose node a on Segment II, has increased range in R^* , but does not have any intended receivers on Segment I. According to Lemma 8, node a must have at least one intended receiver on either Segments III,

IV or V. In this case, we can have another assignment, R , with $R(i) = R^*(i)$, $\forall i \in \mathcal{N} \setminus \{l_{II}, a\}$, $R(a) = d(a, l_{II})$,⁵ and $R(l_{II}) = \max\{R^*(l_{II}), d(l_{II}, r_a^{R^*})\}$, where $r_a^{R^*}$ denotes the farthest intended receiver of node a when it transmits with $R^*(a)$. By using R , the data will be sent to all the nodes with less energy, which contradicts the optimality of R^* . \square

Lemma 13. *In a cross network, utilizing R^* , at most one node in $\hat{\mathcal{N}}$ with increased transmission range exists on Segment II.*

Proof. The proof is by contradiction. Suppose in R^* two nodes in $\hat{\mathcal{N}}$ with increased range exist on Segment II. We denote the node closer to the source by a , and the node closer to the intersection by b . We know that $a \xrightarrow{R^*} b$ (based on Corollary 3), so according to Lemma 11, we have $b \xleftarrow{R^*} a$.

According to Lemma 12, both nodes a and b must have intended receivers on Segment I. Since $a \xleftarrow{R^*} b$, the intended receiver of node b on Segment I is after node $r_{a,I}^{R^*}$ on Segment I (Lemma 7). This results in having the whole transmission circle of node a being inside the transmission circle of node b . To minimize energy (according to Corollary 2) for delivering data from a to b , we can use $a \dashrightarrow r_{a,II}^{R^*}$ (or $a \dashrightarrow b$, if node b is before node $r_{a,II}^{R^*}$), while all the other nodes transmit as before. This way all the nodes receive the data, but node a does not have increased range, which contradicts our assumption. \square

Lemma 14. *For a cross network utilizing the optimal assignment R^* , no more than two nodes with increased range exist on a segment.*

Proof. We first study some properties of the case where in R^* , two nodes from $\hat{\mathcal{N}}$ on one segment have increased transmission range. Then we show that no other node with increased range in R^* can exist on that segment.

Assume nodes a and b have increased range and node a is before node b on the same segment. According to Lemma 8, these two nodes must have some intended receivers on other segments. We know that $a \xleftarrow{R^*} b$ (based on Corollary 3), so according to Lemma 11, we have $b \xleftarrow{R^*} a$.

According to Lemma 7, the intended receivers of node b on any segment must be after the last intended receiver of a on that segment. If node b has any intended receivers

⁵We can have $a \dashrightarrow l_{II}$ to save even more energy. For this, we must have no other nodes with increased range on Segment II, which is proved in Lemma 13.

on the segment aligned to its segment, the whole transmission circle of node a will fall into the range of node b . In this case we just need to deliver the data from node a to node b , and node b transmits the data to all the other-segment receivers of node a . To minimize energy (according to Corollary 2) for delivering data from a to all its same-segment receivers (when it transmits with $R^*(a)$) we can use $a \dashrightarrow r_{a,S_a}^{R^*}$ (or $a \dashrightarrow b$, if node b is before node $r_{a,S_a}^{R^*}$, and b covers the rest), while all the other nodes transmit as before. Note that, according to Corollary 5, the nodes from n_a up to $r_{a,S_a}^{R^*}$ that do not have increased range (i.e., except for node b , if b is before $r_{a,S_a}^{R^*}$), have zero as their transmission range, and have no effect on the delivery of data to other nodes. This means that all the nodes receive the data while node a does not have increased range, which contradicts our assumption. Therefore, the only possible case is that node a has some intended receivers (beyond the range of b) on the aligned segment to its segment, and node b has some intended receivers (beyond the range of a) on (one of the) perpendicular segments to its segment (except for Segment II, according to Lemma 9).

Now, we prove that no other node with increased range exists on segment S_a . We showed in Lemma 13 that no more than one node with increased range exists on Segment II. So, the proof of this lemma for Segment II is already given. Two other cases for the segment of interest remain:

- 1) Segment I.
- 2) One of the Segments III, IV or V.

Case (1): If node a just has intended receivers on Segment II, then node b has an intended receiver on another segment, which results in covering the whole transmission circle of node a , which as we showed before, is not possible. This means that node a has some intended receivers on Segments III, IV, or V. The rest of the proof is by contradiction. Suppose nodes a , b and c , all located on Segment I, have increased range. Without loss of generality, assume that node c is after node b , which is after node a . According to Corollary 3, $a \xrightarrow{R^*} c$ and $b \xrightarrow{R^*} c$. For node c to have increased range, according to Lemma 8, it has to have some intended receivers on other segments rather than S_c . If node c has an intended receiver on one of the other segments, according to Lemma 7, that node is after the last intended receiver of nodes a and b on that segment. Due to the circular shape of transmission range of nodes, if node c has any intended receivers on other segments, all the intended receivers of node b on other segments will fall into the range of node c , and we just need to deliver the data

from node b to all its same-segment receivers. To minimize the energy, according to Corollary 2, we can use $b \dashrightarrow r_{b,S_b}^{R^*}$ (or $b \dashrightarrow c$, if node c is before node $r_{b,S_b}^{R^*}$), while all the other nodes transmit as before, in which node b does not have increased range. This contradicts our assumption.

Case (2): According to Lemma 9, the nodes on Segments III, IV and V do not have intended receivers on Segment II. Using the same approach as in case (1), while ignoring Segment II, the lemma can be proved. \square

Lemma 15. *Consider a cross network that utilizes the optimal assignment R^* . If a node with increased range exists on Segment II, then the nodes on Segments III, IV and V do not have intended receivers on Segment I.*

Proof. The proof is by contradiction. Suppose node a with increased range is located on Segment II, and node b on one of the Segments III, IV or V has intended receivers on Segment I. We know that $a \xleftarrow{R^*} b$ (due to the circular transmission ranges of the nodes), so according to Lemma 7, the intended receivers of node b on Segment I are after node $r_{a,I}^{R^*}$. We can have another assignment, R , with $R(i) = R^*(i)$, $\forall i \in \mathcal{N} \setminus \{l_{II}, a\}$, $R(a) = d(a, l_{II})$,⁶ and $R(l_{II}) = \max\{R^*(l_{II}), d(l_{II}, r_a^{R^*})\}$, where $r_a^{R^*}$ denotes the farthest receiver of node a on Segments III, IV and V, when it transmits with $R^*(a)$. By using R , the data will be sent to all the nodes with less energy, which contradicts the optimality of R^* . \square

Lemma 16. *Consider a cross network that utilizes the optimal assignment R^* . If two nodes with increased range exist, each on one of the Segments I and II, then there is no other node with increased range.*

Proof. Suppose nodes a and b with increased range are located on Segments I and II, respectively. According to Lemma 12, node b has an intended receiver on Segment I. If $a \xleftarrow{R^*} b$, then according to Lemma 7, the intended receivers of b on Segment I are after the last same-segment receiver of a , which, according to Lemma 10, contradicts the optimality of R^* . So we have $a \xleftarrow{R^*} b$.

According to Lemma 7, the intended receivers of node a on any segment are after the last intended receiver of b on that segment. If node b does not have any receivers on Segments III, IV and V, or node a has some intended receivers on either Segments II or III, the whole transmission circle of node b will fall into the range of node a .

⁶We can have $a \dashrightarrow l_{II}$ to save even more energy.

In this case, we just need to deliver the data to node a , and node a transmits the data to all the receivers of node b . Using another transmission range assignment R , with $R(i) = R^*(i)$, $\forall i \in \mathcal{N} \setminus \{s, b\}$, $R(b) = 0$, and $R(s) = \max\{R^*(s), d(s, r_{a,I}^{R^*})\}$, all the nodes receive the data with less energy, which contradicts the optimality of R^* . Therefore, the only possible case is that node b has some intended receivers on Segment III, and node a has some intended receivers (beyond the range of b) on (one of the) Segments IV and V.

Now suppose that we have a third node with increased range (denoted by c). There will be four cases for the location of node c :

- 1) Node c is on Segment I.
- 2) Node c is on Segment II.
- 3) Node c is on Segment III.
- 4) Node c is on either Segments IV or V.

Case (1): Without loss of generality, we assume that node c is after node a . According to Corollary 3, $a \xleftarrow{R^*} c$, and since we have $a \xleftarrow{R^*} b$, so we have $b \xleftarrow{R^*} c$. For node c to have increased range, according to Lemma 8, it has to have some intended receivers on other segments rather than \mathcal{S}_c . If node c has an intended receiver on one of the other segments, according to Lemma 7, that node is after the last intended receiver of nodes a and b on that segment. If node c has any intended receivers on another segment, all the intended receivers of node a on other segments will fall into the range of node c , and we just need to deliver the data from node a to node c . To minimize the energy, according to Corollary 2, we can use $a \dashrightarrow r_{a,I}^{R^*}$ (or $a \dashrightarrow c$, if node c is before node $r_{a,I}^{R^*}$), while other nodes transmit as before. According to Lemma 10, no nodes on Segments II, III, IV, or V have receivers after $r_{a,I}^{R^*}$. Hence node c receives the data from the nodes on its segment. This way, all the nodes receive the data, while node a does not have increased range. This contradicts our assumption. So this case is not possible.

Case (2): In Lemma 13, we showed that this case is not possible.

Before discussing the remaining cases, we prove that for a node d on either Segments IV or V, if we have $d \xleftarrow{R^*} a$, node d does not cover all the receivers of node b on Segment III.

The proof is by contradiction. Suppose for node d , located on one of the Segments IV or V, we have $d \xleftarrow{R^*} a$, and also all the receivers of b on Segment III are in the range of node d . Using another transmission range assignment R , with $R(i) = R^*(i)$,

$\forall i \in \mathcal{N} \setminus \{s, b\}$, $R(b) = 0$, and $R(s) = \max\{R^*(s), d(s, r_{b,I}^{R^*})\}$, all the nodes receive the data⁷ with less energy, which contradicts the optimality of R^* .

Also, note that for any node c on Segments III, IV or V, we have $b \xleftrightarrow{R^*} c$ (Lemma 9). Since the nodes on Segments III, IV or V do not have intended receivers after $r_{b,I}^{R^*}$ on Segment I (Lemma 15), node a does not receive the data via a path containing a node on Segments III, IV or V. So for the remaining two cases, we have $b \xleftrightarrow{R^*} c$ and $a \xleftrightarrow{R^*} c$, which according to Lemma 11, results in having $c \xleftarrow{R^*} b$ and $c \xleftarrow{R^*} a$.

Case (3): We have $c \xleftarrow{R^*} a$. Node a has no intended receivers on Segment III. So, for a node (e.g., c) on Segment III, we have $c \xleftarrow{R^*} a$, if c is after $r_{b,III}^{R^*}$, and a node (e.g., d) on a segment other than Segments I and II exists that $d \xleftarrow{R^*} a$ and d transmits to the nodes after $r_{b,III}^{R^*}$. Otherwise, if c is within the range of b , we will have $c \xleftrightarrow{R^*} a$. Also, if c is after $r_{b,III}^{R^*}$ and node d does not exist, we will have $c \xleftarrow{R^*} e$, where e is a receiver of node b on Segment III, which results in $c \xleftrightarrow{R^*} a$. But we proved that if such node (node d on one of the Segments IV or V, for which we have $d \xleftarrow{R^*} a$) exists, it does not cover all the receivers of node b on Segment III. So this case is not possible.

Case (4): We have $c \xleftarrow{R^*} a$ and $c \xleftarrow{R^*} b$. According to Lemmas 9 and 15 node c has no intended receivers on Segments I and II. The other-segment intended receivers of node c on the two other remaining segments are after the last receivers of nodes a and b on those segments, according to Lemma 7. In both cases, all the receivers of node b on Segment III fall into the range of node c .⁸ We proved that if such a node (node $d = c$ on one of the Segments IV or V, where $(d = c) \xleftarrow{R^*} b$) exists, it does not cover all the receivers of node b on Segment III. So this case is not possible. \square

Proof of Theorem 4. Two cases may happen:

- 1) A node with increased range does not exist on Segment II.
- 2) A node with increased range exists on Segment II.

Case (1.a):

First we consider the case where only one node on any segment (except for Segment II) has increased range in R^* . Suppose there are four nodes with increased range in R^* , and each node is located on one of the Segments I, III, IV and V. Denote these

⁷Node a receives and transmits to all the nodes covered by $R^*(a)$ and $R^*(b)$ on Segments IV and V (including \mathcal{S}_d). Node d receives and transmits to all the nodes covered by $R^*(b)$ on Segment III. The rest of the nodes receive the data similar to when we use R^* .

⁸If c has intended receivers on the perpendicular segment to its segment, we have $\text{cov}_{\perp}^{R^*}(c) > \text{cov}_{oppo}^{R^*}(b)$. If c has intended receivers on the aligned segment to its segment, we have $\text{cov}_{\perp}^{R^*}(c) \geq \text{cov}_{oppo}^{R^*}(c) > \text{cov}_{\perp}^{R^*}(a) > \text{cov}_{\perp}^{R^*}(b) \geq \text{cov}_{oppo}^{R^*}(b)$.

nodes by a , b , c and d . Each of these nodes has to have some intended receivers on other segments than its own segment (Lemma 8). Without loss of generality, and according to Lemma 11, assume $b \xleftarrow{R^*} a$, $c \xleftarrow{R^*} b$ (which results in $c \xleftarrow{R^*} a$) and $d \xleftarrow{R^*} c$ (which means $d \xleftarrow{R^*} a$ and $d \xleftarrow{R^*} b$). So we have $a \nleftrightarrow d$, $b \nleftrightarrow d$ and $c \nleftrightarrow d$. According to Lemmas 7 and 10, node d does not have intended receivers on \mathcal{S}_a , \mathcal{S}_b and \mathcal{S}_c . Since the nodes on Segments III, IV and V do not have intended receivers on Segment II (Lemma 9), the only possible case is that node d is on Segment I and has some intended receivers on Segment II. This means that if node d does not transmit, some nodes on Segment II will not receive the data. Therefore, all the nodes on Segments II, IV and V will not receive the data (circular transmission of the nodes results in reception of data by all the nodes on Segment II before all the nodes on Segments III, IV and V). This contradicts with $d \xleftarrow{R^*} a$, $d \xleftarrow{R^*} b$ and $d \xleftarrow{R^*} c$. Therefore, node d does not have any intended receivers on other segments, which according to Lemma 8 contradicts the optimality of R^* . So, in this case, at most three nodes with increased range exist in R^* .

Case (1.b):

Now assume that in R^* , two nodes from \mathcal{N} on one segment (except for Segment II) have increased transmission range. Assume nodes a and b have increased range and node a is before node b on the same segment. In the proof of Lemma 14, we showed that $b \xleftarrow{R^*} a$. Also, we showed that the only possible case is that node a has some intended receivers on the aligned segment to its segment,⁹ and node b has some intended receivers (beyond the range of a) on (one of the) perpendicular segments to \mathcal{S}_a .

Before further discussion, we prove that for a node d on a segment perpendicular to \mathcal{S}_a , if we have $d \xleftarrow{R^*} b$, node d does not cover all the receivers of node a on a segment aligned to \mathcal{S}_a .

The proof is by contradiction. Suppose for node d , located on a perpendicular segment to \mathcal{S}_a , we have $d \xleftarrow{R^*} b$, and also all the receivers of a on a segment aligned to \mathcal{S}_a are in the range of node d . Using $a \dashrightarrow r_{a,\mathcal{S}_a}^{R^*}$ (or $a \dashrightarrow b$, if node b is before node $r_{a,\mathcal{S}_a}^{R^*}$), while all the other nodes transmit as before, all the nodes in the network receive the data.¹⁰ This way, node a does not have increased range, which is a contradiction.

⁹If $\mathcal{S}_a = I$, this aligned segment is just Segment III (see proof of Lemma 14).

¹⁰Node b receives and transmits to all the nodes covered by $R^*(a)$ and $R^*(b)$ on the perpendicular segments to \mathcal{S}_a (including \mathcal{S}_d). Node d receives and transmits to all the nodes covered by $R^*(a)$ on a segment aligned to \mathcal{S}_a . The rest of the nodes receive the data similar to when we use R^* .

Now suppose that we have a third node with increased range (denoted by c). There will be four cases for the location of node c (except for Segment II):

- 1.b.i) Node c is on the same segment as nodes a and b , i.e., segment \mathcal{S}_a .
- 1.b.ii) Node c is on the segment aligned to the segment of nodes a and b .
- 1.b.iii) Node c is on a segment perpendicular to the segment of nodes a and b , on which node b has intended receivers.
- 1.b.iv) Node c is on a segment perpendicular to the segment of nodes a and b , on which node b does not have any intended receivers.

Case (1.b.i): This case is not possible, according to Lemma 14.

Case (1.b.ii): First we prove that $c \xleftarrow{R^*} a$. Using proof by contradiction, we assume $c \not\xleftarrow{R^*} a$. We know that node a has intended receivers on \mathcal{S}_c . According to Lemma 7, the intended receivers of a on \mathcal{S}_c are after the last same-segment receiver of c , which, according to Lemma 10, contradicts the optimality of R^* . Thus we have $c \xleftarrow{R^*} a$.

Now we prove that $b \xleftarrow{R^*} c$. Suppose node c does not transmit. Since we have $a \xleftarrow{R^*} c$, node a receives the data, and transmits it. If node b is within the range of node a , then b will receive the data as well, and obviously we have $b \xleftarrow{R^*} c$. According to Lemma 10, no nodes on other segments have receivers after $r_{a,\mathcal{S}_a}^{R^*}$. If node b is after $r_{a,\mathcal{S}_a}^{R^*}$, it still receives the data via the nodes on its segment. So, in this case also, we have $b \xleftarrow{R^*} c$.

Since $b \not\xleftarrow{R^*} c$, based on Lemma 11 we have $c \xleftarrow{R^*} b$. This implies node c is not within the range of a . Node b has no intended receivers on \mathcal{S}_c . So, for node c we have $c \xleftarrow{R^*} b$, if c is after $r_{a,\mathcal{S}_c}^{R^*}$, and a node (e.g., d) on a segment other than \mathcal{S}_a and \mathcal{S}_c exists that $d \xleftarrow{R^*} b$ and d transmits to the nodes after $r_{a,\mathcal{S}_c}^{R^*}$. If c is after $r_{a,\mathcal{S}_c}^{R^*}$ and such node d does not exist, we will have $c \xleftarrow{R^*} e$, where e is a receiver of node a on \mathcal{S}_c , which results in $c \not\xleftarrow{R^*} b$. Therefore, node d on one of the perpendicular segments to \mathcal{S}_a and \mathcal{S}_c exists, and we have $d \xleftarrow{R^*} b$. Node d covers all the receivers of node a on segment \mathcal{S}_c . This contradicts what we proved before studying the cases. Therefore, this case is not possible.

Case (1.b.iii): First we prove that $c \xleftarrow{R^*} b$. Using proof by contradiction, we assume $c \not\xleftarrow{R^*} b$. We know that node b has intended receivers on \mathcal{S}_c . According to Lemma 7 the intended receivers of b on \mathcal{S}_c are after the last same-segment receiver of c , which, according to Lemma 10, contradicts the optimality of R^* . So we have $c \xleftarrow{R^*} b$.

Knowing $b \not\xleftarrow{R^*} c$, and using Corollary 4, we have $a \not\xleftarrow{R^*} c$. So, according to Lemma 11, $c \xleftarrow{R^*} a$. If node c has an intended receiver on \mathcal{S}_a , according to Lemma 7, it

transmits to a node after the last same-segment receiver of a , which, according to Lemma 10, contradicts the optimality of R^* . So, node c has no intended receiver on \mathcal{S}_a . The intended receivers of node c on the other remaining segments are after the last receivers of nodes a and b on those segments, according to Lemma 7. This results in having all the receivers of node a on the segments aligned to \mathcal{S}_a to fall into the range of node c .¹¹ Similar to the Case (1.b.ii), if such node (node $d = c$ on one of the perpendicular segments to \mathcal{S}_a , where $(d = c) \xleftarrow{R^*} b$) exists, it does not cover all the receivers of node a on the segments aligned to \mathcal{S}_a . So this case is not possible.

Case (1.b.iv): First we prove that $c \not\xleftrightarrow{R^*} a$. Using proof by contradiction, we assume $c \xleftarrow{R^*} a$, and thus $a \not\xleftrightarrow{R^*} c$. If node c does not transmit, node a receives the data, and transmits it. Using the same approach as in case (1.2), we have $b \xleftarrow{R^*} c$. So, according to Lemma 11, we have $c \xleftarrow{R^*} b$. Similar to the previous case, node c can not have increased range, as it does not have any intended receivers on other segments.

So, Case (1.b.iv) requires $c \not\xleftrightarrow{R^*} a$. According to Lemma 11, we have $a \xleftarrow{R^*} c$, which based on Lemma 5 results in $b \xleftarrow{R^*} c$.

Now, using contradiction, we prove that no fourth node with increased range can exist. Suppose a fourth node with increased range, denoted by d , exists. We showed in the previous cases that node d can not be on \mathcal{S}_a , the segment aligned to \mathcal{S}_a or the segment perpendicular to \mathcal{S}_a on which node b has some intended receivers, while we have two nodes with increased range on \mathcal{S}_a . It cannot be on Segment II as well. Therefore, node d is on the same segment as node c (\mathcal{S}_c). Rename the nodes c and d so that node d be after node c . Similar to the approach we used to prove $c \not\xleftrightarrow{R^*} a$, we prove that $d \not\xleftrightarrow{R^*} a$. Hence according to Lemmas 5 and 11, we have $a \xleftarrow{R^*} d$ and $b \xleftarrow{R^*} d$. Also we know that $c \not\xleftrightarrow{R^*} d$ (Corollary 3). The intended receivers of node d on any segment are after the last intended receiver of c on that segment. Similar to the discussion we had for nodes a and b , here the only possible case is that node d has intended receivers on one the perpendicular segments to \mathcal{S}_d . If we use $c \dashrightarrow r_{c,\mathcal{S}_c}^{R^*}$ (or $c \dashrightarrow d$, if node d is before node $r_{c,\mathcal{S}_c}^{R^*}$), while all the other nodes use the same ranges, all the nodes receive the data¹² with less energy. This contradicts the optimality of

¹¹If c has intended receivers on the perpendicular segment to its segment, we have $\text{cov}_{\perp}^{R^*}(c) > \text{cov}_{oppo}^{R^*}(a)$. If c has intended receivers on the aligned segment to its segment, we have $\text{cov}_{\perp}^{R^*}(c) \geq \text{cov}_{oppo}^{R^*}(c) > \text{cov}_{\perp}^{R^*}(b) > \text{cov}_{\perp}^{R^*}(a) \geq \text{cov}_{oppo}^{R^*}(a)$.

¹²All the receivers of c on the perpendicular segments to \mathcal{S}_c are covered by d . The receivers of c on the aligned segments to \mathcal{S}_c are covered by d (if the aligned segment is Segment II) or b (which receives the data from d). That's because according to Lemma 7, the intended receivers of node b on those segments are after the last intended receiver of c on them.

R^* .

Case (2): If a node with increased range exists on Segment I, then according to Lemma 16, no more nodes with increased range exist. For the rest of the proof, we assume that no node with increased range exists on Segment I.

First, consider the case where only one node on each of the Segments III, IV and V has increased range in R^* . The proof is by contradiction. Suppose there are four nodes with increased range in R^* , and each node is located on one of the Segments II, III, IV and V. Denote these nodes by a , b , c and d , where node a is located on Segment II. According to Lemmas 9 and 15, nodes b , c and d do not have intended receivers on Segments I and II. Without loss of generality, and according to Lemma 11, assume $c \xleftarrow{R^*} b$, $d \xleftarrow{R^*} c$ (which results in $d \xleftarrow{R^*} b$). So we have $b \xleftarrow{R^*} d$ and $c \xleftarrow{R^*} d$. According to Lemmas 7 and 10, node d does not have intended receivers on \mathcal{S}_b and \mathcal{S}_c . Therefore, node d does not have any intended receivers on other segments, which according to Lemma 8 contradicts the optimality of R^* . So, in this case, at most three nodes with increased range exist in R^* .

Now, we consider the case where more than one node with increased range may exist on each of the Segments III, IV and V. According to Lemma 14, no more than two nodes with increased range exist on each segment. Also, according to Lemma 13, we have at most one node with increased range on Segment II. So, assume nodes a and b on one of the Segments III, IV or V have increased range, while we have a node with increased range on Segment II. We assume that node a is before b . Hence, according to Corollary 3 and Lemma 11, we have $b \xleftarrow{R^*} a$. If these two nodes are on Segment III, as they can not have intended receivers on Segments I and II (Lemmas 9 and 15), they both have intended receivers on Segments IV and V. Node b having intended receivers on (one of the) Segments IV and V results in having the whole transmission circle of a within range b , which means that node a does not need to have increased range. This contradicts our assumption, so this case can not happen. If nodes a and b are on one of the Segments IV or V (e.g., Segment IV), as they can not have intended receivers on Segments I and II (Lemmas 9 and 15), they have intended receivers on Segments III and V. Similar to the approach used in cases (1.1), (1.2) and (1.3), we can prove that no other node with increased range exists. So, in this case, at most three nodes with increased range exist in R^* . \square

To find R^* , one has to search among all the possible assignments, constructed from all the possible combinations of the nodes with increased transmission range and the

values of their transmission ranges. Based on Theorem 4, one needs to consider all the possible combinations of three nodes chosen from $\hat{\mathcal{N}}$, i.e., $\binom{N-5}{3}$ choices, and for each such node, all the N possible range values including zero. For the nodes in the set $\{l_{II}, f_{III}, f_{IV}, f_V\}$ too, each node can take any of the N possible range values. For the source node s , all the non-zero range values can be selected. To describe the eight nodes with possible increased range, we first define set \mathcal{T} to contain all the possible choices for the three nodes from set $\hat{\mathcal{N}}$, and concatenate each of the possible choices for the three nodes with the nodes l_{II}, f_{III}, f_{IV} , and f_V . Therefore, each member $t \in \mathcal{T}$ is denoted by $\{t_1, t_2, t_3, t_4, t_5, t_6, t_7\}$, where t_1, t_2 and t_3 represent the three chosen nodes, and t_4, t_5, t_6 , and t_7 denote the nodes l_{II}, f_{III}, f_{IV} , and f_V , respectively. Set \mathcal{T} has (at most) $\binom{N-5}{3}$ members. To search all the possible ranges for each of the three selected nodes from $\hat{\mathcal{N}}$, and also the nodes in the set $\{s, l_{II}, f_{III}, f_{IV}, f_V\}$, we construct the set $\mathcal{C} = \mathcal{N}^8$. Each member of \mathcal{C} , which is denoted by \mathbf{c} , is considered as an 8-tuple of form $(c_0, c_1, c_2, c_3, c_4, c_5, c_6, c_7)$. The range of the source node is equal to $d(s, c_0)$, and the range of node t_j (for $1 \leq j \leq 7$) is equal to $d(t_j, c_j)$.

To account for the order in which nodes on different segments transmit data, we also need to search among all the $5! = 120$ different segment orderings. We define set \mathcal{P} to contain all the possible permutations of the Segments $\{I, II, III, IV, V\}$. Each segment ordering $\mathbf{p} = (p_1, p_2, p_3, p_4, p_5) \in \mathcal{P}$ contains the labels of the segments in the order they have to be checked. On each segment, we assign ranges to the nodes starting from the first node up to the last node on that segment. For the source node and the seven nodes from set \mathcal{T} , the range is already assigned as discussed before. For each remaining node, we assign its M value to the range, if the next adjacent neighbor of the node has not received the data yet, otherwise, zero is assigned as the range (based on Corollary 5).

Since all possible assignments do not result in the delivery of data to all the nodes, we need to construct the desired assignments in a way that the delivery of data to all the nodes is guaranteed. This is done by using received labels (label \mathbf{r}). If a node receives the data, we tag it by label \mathbf{r} . An assignment in which a node without this label exists, will be ignored. The optimal assignment is the assignment among all the constructed assignments which has the minimum cost.

To reduce the time complexity of the algorithm, we define N^2 sets $\mathcal{R}_{i,j}, \forall i, j \in \mathcal{N}$. Each set $\mathcal{R}_{i,j}$ contains all the nodes that are within the transmission range of node i , when $R(i) = d(i, j)$. Construction of each of these sets can be done in $\mathcal{O}(N)$ time.

Hence, all these sets can be acquired with time complexity $\mathcal{O}(N^3)$. Note that the construction of these sets is performed before the execution of the algorithm, and thus does not introduce any additional time complexity to the algorithm.

The pseudo code of the algorithm that finds the optimal transmission range assignment (i.e., R^*) is given as Algorithm 1: Optimal Range Assignment. In this algorithm, by constructing sets \mathcal{P} , \mathcal{T} , \mathcal{C} and $\mathcal{R}_{i,j}, \forall i, j \in \mathcal{N}$, we search for the at most three nodes with increased range. In the process of finding the nodes with increased range, we assign ranges to all the nodes of the network. We then choose the assignment with the minimum energy consumption as the optimal assignment.

Algorithm 1 Optimal Range Assignment

Input: exact location of all the nodes

Output: R^* (optimal transmission range assignment)

```

1: assign  $cost^* = +\infty$ 
2: construct sets  $\mathcal{P}$ ,  $\mathcal{T}$ ,  $\mathcal{C}$  and  $\mathcal{R}_{i,j}, \forall i, j \in \mathcal{N}$ 
3: for all  $t \in \mathcal{T}$  do
4:   for all  $c \in \mathcal{C}$  do
5:     for all  $p \in \mathcal{P}$  do
6:       clear transmission range assignment  $R$ 
7:       erase the  $\tau$  labels of all the nodes
8:       tag node  $s$  with label  $\tau$ 
9:       assign  $R(s) = d(s, c_0)$ 
10:      calculate  $cost(R) = R^\alpha(s)$ 
11:      tag all the nodes in set  $\mathcal{R}_{s,c_0}$  with label  $\tau$ 
12:      for  $k = 1$  to 5 do
13:        for all nodes  $n$  on segment  $p_k$  (from the first node to the last node) do
14:          if node  $n$  has not received the data yet (i.e., does not have label  $\tau$ ) then
15:            stop and go to next  $p$ 
16:          else if  $n = t_j$  for  $1 \leq j \leq 7$  then
17:            assign  $R(n) = d(n, c_j)$ 
18:            tag all the nodes in set  $\mathcal{R}_{n,c_j}$  with label  $\tau$ 
19:          else if node  $n_n$  exists and does not have label  $\tau$  then
20:            assign  $R(n) = M(n)$ 
21:            tag all the nodes in set  $\mathcal{R}_{n,n_n}$  with label  $\tau$ 
22:          else
23:            assign  $R(n) = 0$ 
24:          end if
25:          update  $cost(R) = cost(R) + R^\alpha(n)$ 
26:        end for
27:      end for
28:      if all the nodes receive the data and  $cost(R) < cost^*$  then
29:        assign  $cost^* = cost(R)$  and  $R^* = R$ 
30:      end if
```

```

31:   end for
32: end for
33: end for

```

Theorem 5. *Algorithm 1: Optimal Range Assignment, finds R^* with time complexity $\mathcal{O}(N^{12})$.*

Proof. There are $\mathcal{O}(N^{11})$ choices for the selection of nodes in the set \mathcal{T} , and range assignments from \mathcal{C} . For each such choice, there are $N - 8$ remaining nodes whose range is assigned as either zero or their M value. \square

A Lower Bound on the Optimal Energy Consumption

In this section, we add a node, denoted by o , to the intersection of the cross network, and then find the optimal transmission range assignment of the new network. In the following theorem, we prove that the optimal solution of this new network can be used to obtain a lower bound on the optimal energy consumption of the original network.

Theorem 6. *The optimal transmission range assignment of a cross network, with at least one node on either Segments III, IV or V, in which a node is added to the intersection has cost smaller than or equal to the cost of the optimal transmission range assignment of the original network.*

Before proving the above theorem, we state the following definition and theorem, which are valid for any wireless network.

Definition 9. *Consider a wireless network (of any dimension) in which a source node broadcasts the data to all the other nodes. Furthermore, consider the set of **all** possible transmission range assignments that guarantee delivery of data to all nodes in the network. The **covered space** of that network consists of the union of all subspaces in which a potential receiver will receive the data for any of these transmission range assignments.*

Note that no restriction on the shape of the coverage area of the nodes is assumed in the above definition. As long as the space is covered by *all* the transmission range assignments that guarantee delivery of data to all nodes in the network, the definition and the following theorem are valid.

Theorem 7. *If some nodes are added to the covered space of a wireless network of any dimension (in which a source node broadcasts the data to all the other nodes), the cost of the optimal transmission range assignment of the new network is less than or equal to that of the original network (without the added nodes).*

Proof. According to the definition of the covered space of a network, for the new network (with the added nodes), all the range assignments of the original network (defined for the original nodes of the network) deliver the data not only to all the original nodes, but also to the added nodes as well. Adding the new nodes will result in being able to have some additional range assignments to deliver the data to all the (original and new) nodes of the network. So the set of all the range assignments of the new network contains all the range assignments of the original network plus those new assignments. The optimal transmission range assignment of the new network has the minimum cost among all the range assignments, which means it is smaller than or equal to the cost of the optimal range assignment of the original network. \square

Finding the covered space of the networks is not a simple task, but for some cases, like adding a node to the intersection of a cross network (with at least one node on either Segments III, IV or V), it is easy to see that the node is actually added to a point which belongs to the covered space of the network.

Proof of Theorem 6. Since we assume circular transmission range for the nodes, delivering the data to the node(s) located on Segments III, IV and V results in the delivery of the data to a potential node located at the intersection. Hence, for all the assignments of a cross network (with at least one node on either Segments III, IV or V), a node located at the intersection is in the covered space of the network. The proof then follows from Theorem 7. \square

Theorem 8. *Finding the optimal transmission range assignment of the network with an added node at the intersection of a cross network has time complexity $\mathcal{O}(N^9)$.*

Proof. For the new networks, we denote the set of nodes, containing all the nodes from the original network plus the added node o , as $\mathcal{N}^{new} = \mathcal{N} \cup \{o\}$. We substitute the sets \mathcal{N} and $\hat{\mathcal{N}}$ by sets \mathcal{N}^{new} and $\mathcal{N}^{new} \setminus \{s, o\}$, respectively. We also replace all the nodes in the set $\{l_{II}, f_{III}, f_{IV}, f_V\}$ by the added node o . By making these changes, we can easily see that all the lemmas, corollaries and Theorem 4 presented in Section 4.2.1 are valid for the new network. In this new setting, since four nodes each with

N possible ranges are substituted by only one such node, the search space will be reduced by a factor of N^3 . This results in having a time complexity $\mathcal{O}(N^9)$. \square

4.2.2 Near-Optimal Range Assignment With Linear Time Complexity

The focus of the proposed sub-optimal algorithm is to find the nodes that can save energy by not transmitting, whereas the other nodes only transmit at a power level that is needed for their next adjacent neighbors to receive the data. As we will find out later, this algorithm performs close to optimal. In the following, we thus refer to it as being “near-optimal.”

Similar to the optimal algorithm, and to check all the segment orderings, we define set \mathcal{P} to contain all the 120 possible permutations of the Segments {I,II,III,IV,V}. For any node n on each of the ordered segments, we assign $R(n) = M(n)$, if node n has not received the data through the previously assigned transmission ranges. Otherwise, we assign $R(n) = 0$. Here, unlike the optimal algorithm, and to reduce time complexity, we do not construct sets $\mathcal{R}_{i,j}$. Instead, we just find the receivers of a fixed number of nodes (not a function of N). This has time complexity $\mathcal{O}(N)$.

To guarantee the delivery of data to all the nodes, we check the first node of Segments III, IV and V to see if it receives the data or not. If not, we change the range of node s ¹³ in the set $\{s, l_{\text{II}}, f_{\text{III}}, f_{\text{IV}}, f_{\text{V}}\}$ so that it delivers the data to the first node of the segment of interest. If Segment II is empty, we assume that node s takes all the functionalities of node l_{II} .

The near-optimal algorithm is described in Algorithm 2: Near-Optimal Linear-Time Algorithm. It has output R^{NO} as the near-optimal range assignment. In this algorithm we assign the $M(\cdot)$ values to the nodes. We then reassign the transmission range of the nodes with unnecessary transmission back to zero, to save energy.

Algorithm 2 Near-Optimal Linear-Time Algorithm

Input: exact location of all the nodes

Output: R^{NO} (near-optimal transmission range assignment)

- 1: assign $cost^{NO} = +\infty$
- 2: construct set \mathcal{P}
- 3: **for all** $p \in \mathcal{P}$ **do**
- 4: clear transmission range assignment R
- 5: erase the τ labels of all the nodes

¹³See line 30 of Algorithm 2: Near-Optimal Linear-Time Algorithm, for definition.

```

6:  tag node  $s$  with label  $\tau$ 
7:  assign  $R(s) = d(s, f_{p_1})$ , where  $f_{p_1}$  denotes the first node on segment  $p_1$ 
8:  calculate  $cost(R) = R^\alpha(s)$ 
9:  tag all the receivers of node  $s$  with label  $\tau$ 
10: for  $k = 1$  to 5 do
11:   for all nodes  $n$  on segment  $p_k$  (from the first node to the last node) do
12:     if node  $n$  has not received the data yet (i.e., is not tagged by label  $\tau$ ) then
13:       stop and go to next p
14:     else if node  $n$  exists and does not have label  $\tau$  then
15:       assign  $R(n) = M(n)$ 
16:       tag node  $n$  with label  $\tau$ 
17:     else
18:       assign  $R(n) = 0$ 
19:     end if
20:     update  $cost(R) = cost(R) + R^\alpha(n)$ 
21:   end for
22:   find  $m_\perp^{p_k} = \arg \max_{m \text{ on segment } p_k} \{cov_\perp^R(m)\}$ ,  $m_{oppo}^{p_k} = \arg \max_{m \text{ on segment } p_k} \{cov_{oppo}^R(m)\}$ 
23:   if  $p_k = I$  then
24:     find  $m_{II}^I = \arg \max_{m \text{ on Segment I}} \{R(m) - d(s, m)\}$ 14
25:   else if  $p_k = II$  then
26:     find  $m_I^{II} = \arg \max_{m \text{ on Segment II}} \{R(m) - d(s, m)\}$ 15
27:   end if
28:   tag all the nodes that receive the data from nodes  $m_\perp^{p_k}$ ,  $m_{oppo}^{p_k}$ ,  $m_{II}^I$ , or  $m_I^{II}$  (if they exist) with label  $\tau$ 
29:   if  $p_k \neq I$  and  $p_{k+1} \in \{III, IV, V\}$  and node  $f_{p_{k+1}}$  has not received the data yet then
30:     reassign  $R(sn) = d(sn, f_{p_{k+1}})$ , where  $sn = \begin{cases} l_{II}, & \text{if } p_k = II \\ s, & \text{if } p_k = II, \text{ and Segment II is empty} \\ f_{p_k}, & \text{otherwise} \end{cases}$ 
31:     tag all the nodes that receive the data from node  $sn$  with label  $\tau$ 
32:     if  $p_k \in \{III, IV, V\}$  then
33:       reassign  $R(i) = 0$  for all nodes  $i \neq sn$  on segment  $p_k$ , for which node  $n_i$  is a receiver of node  $sn$ 
34:     end if
35:   end if
36: end for
37: if all the nodes receive the data and  $cost(R) < cost^{NO}$  then

```

¹⁴Node m_{II}^I denotes the node among all the nodes on Segment I, that has the maximum coverage on Segment II.

¹⁵Node m_I^{II} denotes the node among all the nodes on Segment II, that has the maximum coverage on Segment I.

```

38:      assign  $cost^{NO} = cost(R)$  and label  $R^{NO} = R$ 
39:  end if
40: end for

```

4.2.3 Distributed Range Assignment

In the distributed algorithm, every node in set $\hat{\mathcal{N}}$ just needs to know the distance to its next adjacent neighbor. In this algorithm, the source node transmits the data to its two adjacent neighbors in Segments I and II (we will discuss the case where Segment II is empty later), by having a transmission range sufficient to reach the farthest one. Any node $a \in \hat{\mathcal{N}}$ waits to receive the data for the first time, then transmits the data to its next adjacent neighbor, i.e., with range $M(a)$.

If $a \in \{l_{II}, f_{III}, f_{IV}, f_V\}$, in addition to its next adjacent neighbor, node a has to consider the other three nodes in this set as well. The graph consisting of the nodes in the set $\{l_{II}, f_{III}, f_{IV}, f_V\}$, and all the possible edges between them, is denoted by G^\diamondsuit , and is illustrated in Fig. 7.2.¹⁶ We refer to this graph as the *diamond graph*.

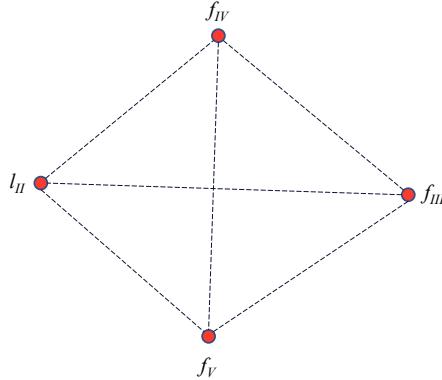


Figure 4.4: The graph G^\diamondsuit , for which its MST is needed to be found in each of the nodes in the set $\{l_{II}, f_{III}, f_{IV}, f_V\}$.

All the nodes in the set $\{l_{II}, f_{III}, f_{IV}, f_V\}$ calculate the MST of the graph G^\diamondsuit , e.g., using Prim's algorithm. The transmission range of each of the nodes in the set $\{l_{II}, f_{III}, f_{IV}, f_V\}$ is then selected as the largest of the maximum length of the node's connected edges in the MST of G^\diamondsuit and its distance to its next adjacent neighbor. If Segment II is empty, node s replaces node l_{II} and its range is the largest of the

¹⁶If Segment II is empty, node l_{II} will be replaced by node s .

maximum length of its connected edges in the MST of G^\diamond and its distance to node f_I .

The distributed range assignment, denoted by R^D , is thus given by:

$$R^D(s) = \begin{cases} \max\{d(s, f_I), d(s, f_{II})\}, & \text{if Segment II is not empty;} \\ \max\{d(s, f_I), \max_{u:(s,u) \in MST(G^\diamond)}\{d(s, u)\}\}, & \text{if Segment II is empty,} \end{cases} \quad (4.6)$$

$$R^D(a) = \max\{M(a), \max_{u:(a,u) \in MST(G^\diamond)}\{d(a, u)\}\}, \quad \text{for } a \in \{l_{II}, f_{III}, f_{IV}, f_V\},$$

$$R^D(a) = M(a), \quad \text{otherwise.}$$

It is easy to see that the distributed range assignment results in all the network nodes receiving the data. The following theorem shows that the proposed distributed algorithm, with time complexity only $\mathcal{O}(1)$, results in the same range assignment as the algorithm of [41], with time complexity $\mathcal{O}(N^2)$.

Theorem 9. *The distributed transmission range assignment R^D , is the same as the transmission range assignment of the MST-based algorithm of [41] for the cross network.*

Proof. First we find the MST of the graph corresponding to a cross network. The graph consists of the nodes, as its vertices, and there exists an edge between any two nodes of the network, with weight equal to the distance between them. We assume that the weights of the edges are different, and the MST is unique. In the following, we use the cut¹⁷ property of MST, which states that for any cut in the graph, if the weight of an edge crossing the two partitions of the cut is strictly smaller than the weights of all other crossing edges, then this edge belongs to the MST of the graph (for more detail, see, e.g., Theorem 23.1 of [69]). A direct corollary of the cut property (described as Corollary 23.2 in [69]) is that if we partition a graph into disjoint sets of vertices, the edge with the minimum weight among all the crossing edges (in the original graph) between any two disjoint partitions belongs to the MST of the graph. So one way to construct the MST of a graph is to partition it into disjoint vertex

¹⁷A cut $[\mathcal{V}, \bar{\mathcal{V}}]$ is a partition of the vertices of a graph into two disjoint subsets \mathcal{V} and $\bar{\mathcal{V}}$ that are joined by at least one edge [68].

sets, find the MST of the subgraph induced by each set, and link the MST's via the minimum weight crossing edge between them.

We partition the network graph into seven (or five, if Segment II is empty) partitions. The set of partitions consists of the source node alone as one partition, graph G^\diamondsuit as another partition, and five (or four) other partitions each consisting of just the remaining nodes of each segment (i.e., one partition for each segment). If Segment II is empty, node s will be a node in G^\diamondsuit , and there will be only five partitions.

The MST of the s only partition is s itself, and we already discussed the MST of the G^\diamondsuit partition. For any node $a \in \hat{\mathcal{N}}$, we define set \mathcal{V}_a to contain node a and all the nodes after that on segment \mathcal{S}_a . Set $\bar{\mathcal{V}}_a$ contains all the nodes (in set $\hat{\mathcal{N}}$) on segment \mathcal{S}_a , that are not in set \mathcal{V}_a . Consider cut $[\mathcal{V}_a, \bar{\mathcal{V}}_a]$ for the subgraph of segment \mathcal{S}_a . Since all the vertices of this subgraph are on a straight line, the minimum weight (distance) edge crossing $[\mathcal{V}_a, \bar{\mathcal{V}}_a]$ is between nodes a and b , where $\mathbf{n}_b = a$. This edge, according to the cut property of MST, belongs to the MST of the network graph. Using the same approach for all the cuts of a subgraph, we find all the connecting edges, which all together construct a spanning tree of the subgraph. All these edges have to be included in the MST of each subgraph, and no other edge is necessary for forming a spanning tree. By connecting the different partitions of the graph using the minimum weight (distance) edge between any two partitions the MST of the network graph is obtained. The connecting edge of partition of Segment I and partition of s is the edge between s and f_I . Also, the edge between s and f_{II} is the edge between partition s and partition of Segment II (if Segment II is not empty). The three remaining segment partitions are connected to the partition of G^\diamondsuit by the edges between the first nodes of each segment (members of partition G^\diamondsuit) and the second nodes of the same segments (each member of a different partition). We root the tree at s .

In the MST-based transmission range assignment, every node transmits with the range equal to the maximum edge weight (distance) to its children. By observing the way that the MST of the network graph is constructed, we can see that the MST-based transmission range assignment is exactly the same as the distributed transmission range assignment. \square

4.3 Special Case: Source at the Intersection

In this section, we consider cross networks in which the source node is located at the intersection of the two lines.¹⁸ In such networks, we assume that Segment II still exists, but is empty. Hence, node s takes all the functionalities of node l_{II} . Furthermore, all the nodes in the set $\{l_{II}, f_{III}, f_{IV}, f_V\}$ are replaced by the source node s . By doing this, we can see that all the lemmas, corollaries and Theorem 4 presented in Section 4.2.1 are valid for this special-case network. Since, for this network, five nodes ($s, l_{II}, f_{III}, f_{IV}, f_V$), each with N possible ranges are substituted by only one such node (s), the search space will be reduced by a factor of N^4 . This results in having the time complexity of $\mathcal{O}(N^8)$.

4.4 A More General Case: Grid Networks

We can use the proposed distributed algorithm to find a cost-efficient transmission range assignment for grid networks with perpendicular line-segments. The structure of the grid can be arbitrary, e.g., the one shown in Fig. 7.3.

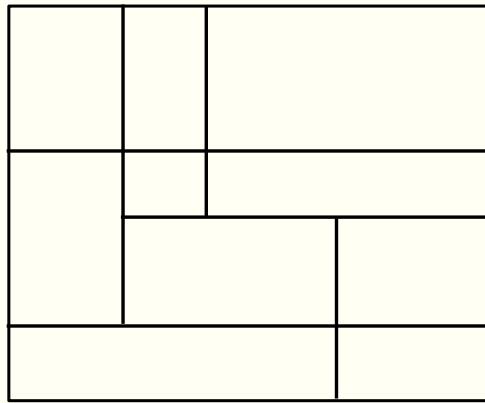


Figure 4.5: A possible structure for grid networks with perpendicular line-segments.

To apply the proposed distributed algorithm, we assume that there is at least one node on each line-segment of the network. This allows for the construction of diamond graphs for each intersection of the network. One can then run the distributed

¹⁸The work in this section is described with more details in [70].

algorithm on each intersection of the grid network similar to what was done for cross networks. There may exist multiple diamond graphs, and the MST of each of them has to be obtained. Note that this can be done in parallel, as there is no dependency between the different diamond graphs and their MST's.

In Section 4.6, we compare the distributed algorithm to BIP (with and without the sweeping procedure), for a two by two square grid network, and show that the performance difference between the two algorithms is rather small, even though our algorithm is distributed and with time complexity $\mathcal{O}(1)$, while BIP is centralized and with time complexity $\mathcal{O}(N^2)$.

4.5 Asymptotic Analysis

In this section, we study the properties of the cross networks in the asymptotic regime where the number of network nodes tends to infinity. This analysis is used to show that when a large number of nodes are distributed uniformly in the network, the difference between the cost of different algorithms is small. We perform the analysis in two steps.

Step 1:

We prove that for cross networks, in which the distance between any two nodes on a line of the network is less than a constant value δ , the difference between the cost of the worst range assignment, i.e., the distributed assignment, and the best one, i.e., the optimal assignment, is upper bounded by a linear function of δ^α . Furthermore, we show that the nodes with increased range are in a certain vicinity of the source, or the intersection.

Step 2:

We prove that cross networks with uniformly distributed nodes on any fixed total length L , can exist, in which the distance between any two nodes on a line of the network is less than any constant value $\delta \leq L$ with probability $1 - \epsilon$, where ϵ depends on the number of nodes in the network, i.e., N . Furthermore, we show that such networks exist for any (fixed) values of L and δ asymptotically almost surely (a.a.s. for short).¹⁹

¹⁹This means that for any (fixed) value of L and $\delta \leq L$, when N goes to infinity, the probability $1 - \epsilon$ converges to one.

4.5.1 Step 1:

We consider cross networks where the distance between any two nodes on a segment is less than a constant value δ . This means that there exists no empty interval with length δ or more on each of the perpendicular lines. Hence, there exist at least $\lceil \frac{d(n_1, n_2)}{\delta} \rceil - 1$ ²⁰ nodes, between any two nodes n_1 and n_2 . The data can be transmitted from node n_1 to node n_2 via the nodes between them ($n_1 \dashrightarrow n_2$), in at most $\lceil \frac{d(n_1, n_2)}{\delta} \rceil$ hops. The length of each hop is at most δ . Thus the cost of sending data by $n_1 \dashrightarrow n_2$ is at most $\lceil \frac{d(n_1, n_2)}{\delta} \rceil \delta^\alpha$ (based on Corollary 2). Using this result, we present the following theorems.

Theorem 10. *For a cross network, if the distance between any two nodes on each line of the network is less than a constant value δ , the difference between $\text{cost}(R^D)$ and $\text{cost}(R^*)$ is less than $\mathcal{K}\delta^\alpha$, where \mathcal{K} is a constant less than 128, in general, and less than 36 for the special case where the source node is located at the intersection.*

Proof. We denote the difference between $\text{cost}(R^D)$ and $\text{cost}(R^*)$, caused by node a with increased range, or the source node, by \mathcal{D}_a . When node a transmits the data with $R^*(a)$, all the receivers of a will receive the data with cost $(R^*(a))^\alpha$. On the other hand, delivering the data to all those nodes (i.e., the receivers of a in the optimal assignment), using the distributed assignment results in a different energy consumption, which we denote by Cost_a^D . So, we have $\mathcal{D}_a = \text{Cost}_a^D - (R^*(a))^\alpha$.

We consider the following four cases for node a to calculate \mathcal{D}_a .

- 1) Node a is the source node.
- 2) Node a is located on Segment I.
- 3) Node a is located on Segment II.
- 4) Node a is located on one of the Segments III, IV or V.

We use the following inequality for the optimal assignment throughout the proof:

$$R^*(a) > (\lceil \frac{R^*(a)}{\delta} \rceil - 1)\delta. \quad (4.7)$$

Case (1): We have two sub-cases:

- 1.1) All the receivers of s when it transmits with $R^*(s)$ are on Segments I and II.
- 1.2) Node s has some receivers on Segments III, IV or V, when it transmits with $R^*(s)$.

²⁰Notation $\lceil r \rceil$ denotes the smallest integer greater than or equal to r .

Case (1.1): For the distributed approach, node s transmits to nodes f_I and f_{II} with range at most δ . Then we have $f_I \dashrightarrow r_{s,I}^{R^*}$ and $f_{II} \dashrightarrow r_{s,II}^{R^*}$. Knowing that $d(f_S, r_{s,S}^{R^*}) \leq R^*(s)$, $\forall S \in \{I, II\}$, we have:

$$Cost_s^D \leq (2 \lceil \frac{R^*(s)}{\delta} \rceil + 1) \delta^\alpha, \quad (4.8)$$

and thus

$$\begin{aligned} \mathcal{D}_s &< (2 \lceil \frac{R^*(s)}{\delta} \rceil + 1) \delta^\alpha - ((\lceil \frac{R^*(s)}{\delta} \rceil - 1) \delta)^\alpha \\ \stackrel{\alpha \geq 2}{\Rightarrow} \quad \mathcal{D}_s &< 4\delta^\alpha. \end{aligned} \quad (4.9)$$

Case (1.2): Based on the assumption of having no empty δ -length section, node l_{II} is located in the δ -neighborhood of the intersection. In the distributed assignment, node s transmits with range at most δ to reach its neighbors. We have $f_{II} \dashrightarrow l_{II}$, then node l_{II} transmits to nodes f_{III} , f_{IV} , and f_V with range at most $\sqrt{2}\delta$, and $f_S \dashrightarrow r_{s,S}^{R^*}$, $\forall S \in \{I, III, IV, V\}$.

We know that $d(f_S, r_{s,S}^{R^*}) \leq R^*(s)$, $\forall S \in \{I, IV, V\}$. Also, $d(f_{II}, l_{II}) + d(f_{III}, r_{s,III}^{R^*}) \leq R^*(s)$. Hence, we have:

$$Cost_s^D \leq (4 \lceil \frac{R^*(s)}{\delta} \rceil + 1 + (\sqrt{2})^\alpha) \delta^\alpha. \quad (4.10)$$

This results in:

$$\begin{aligned} \mathcal{D}_s &< (4 \lceil \frac{R^*(s)}{\delta} \rceil + 1 + (\sqrt{2})^\alpha) \delta^\alpha - ((\lceil \frac{R^*(s)}{\delta} \rceil - 1) \delta)^\alpha \\ \stackrel{2 \leq \alpha \leq 6}{\Rightarrow} \quad \mathcal{D}_s &< 16\delta^\alpha. \end{aligned} \quad (4.11)$$

Case (2): Similar to the previous case, we consider the following two sub-cases:

- 2.1) All the receivers of a when it transmits with $R^*(a)$ are on Segments I and II.
- 2.2) Node a has some receivers on Segments III, IV or V when it transmits with $R^*(a)$.

Case (2.1): We have:

$$\begin{aligned} \mathcal{D}_a &< (2\lceil \frac{R^*(a)}{\delta} \rceil + 1)\delta^\alpha - ((\lceil \frac{R^*(a)}{\delta} \rceil - 1)\delta)^\alpha \\ \stackrel{\alpha \geq 2}{\Rightarrow} \quad \mathcal{D}_a &< 4\delta^\alpha. \end{aligned} \quad (4.12)$$

Case (2.2): In the distributed algorithm the data is sent from f_{II} to l_{II} . Having in mind that $d(f_{II}, l_{II}) + d(f_{III}, r_{a,III}^{R^*}) \leq R^*(a)$, we have:

$$\begin{aligned} \mathcal{D}_a &< (4\lceil \frac{R^*(a)}{\delta} \rceil + 1 + (\sqrt{2})^\alpha)\delta^\alpha - ((\lceil \frac{R^*(a)}{\delta} \rceil - 1)\delta)^\alpha \\ \stackrel{2 \leq \alpha \leq 6}{\Rightarrow} \quad \mathcal{D}_a &< 16\delta^\alpha. \end{aligned} \quad (4.13)$$

Case (3): Here, node a can be a node with increased range, or node l_{II} , which covers some nodes on other segments. Similarly, we have two sub-cases:

- 3.1) All the receivers of a when it transmits with $R^*(a)$ are on Segments I and II.
- 3.2) Node a has some receivers on Segments I, III, IV or V, when it transmits with $R^*(a)$. Note that according to Lemma 12, in the optimal assignment, node a must have some receivers on Segment I.

We can see that for case (3.1), Equation (4.12), and for case (3.2), Equation (4.13) is valid.

Case (4): Node a can be a node with increased range, or node f_{S_a} , and it covers some nodes on other segments. Following the same approach as in previous cases, one can show that Inequality (4.13) holds for this case as well.

The maximum difference between the two costs $\text{cost}(R^D)$ and $\text{cost}(R^*)$ happens if there are three nodes with increased range, and all the nodes in the set $\{s, l_{II}, f_{III}, f_{IV}, f_V\}$ cover more than one node. By considering the worst case for each of these eight nodes, and ignoring the effects that the transmission of data by each of them has on the receivers of the other nodes, we can obtain the following rather loose upper bound on $\text{cost}(R^D) - \text{cost}(R^*)$.

$$\text{cost}(R^D) - \text{cost}(R^*) < \mathcal{K}\delta^\alpha, \quad (4.14)$$

where $\mathcal{K} \leq 128$.

For the case where the source node is located at the intersection, the following two

cases for node a are considered to calculate \mathcal{D}_a .

- 1) Node a is the source node.
- 2) Node a is not the source node.

Case (1): For the distributed approach, node s transmits to nodes $f_{\mathcal{S}}$, $\forall \mathcal{S} \in \{\text{I,II,III,IV}\}$ with range at most δ . Then we have $f_{\mathcal{S}} \dashrightarrow r_{s,\mathcal{S}}^{R^*}$, $\forall \mathcal{S} \in \{\text{I,II,III,IV}\}$. Obviously, $d(f_{\mathcal{S}}, r_{s,\mathcal{S}}^{R^*}) \leq R^*(s)$, $\forall \mathcal{S} \in \{\text{I,II,III,IV}\}$. Therefore:

$$Cost_s^D \leq (4 \lceil \frac{R^*(s)}{\delta} \rceil + 1) \delta^\alpha. \quad (4.15)$$

Using Equations (4.7) and (4.15), we have:

$$\begin{aligned} \mathcal{D}_s &< (4 \lceil \frac{R^*(s)}{\delta} \rceil + 1) \delta^\alpha - ((\lceil \frac{R^*(s)}{\delta} \rceil - 1) \delta)^\alpha \\ \xrightarrow{\alpha \geq 2} \quad \mathcal{D}_s &< 9\delta^\alpha. \end{aligned} \quad (4.16)$$

Case (2): In the distributed approach, we have $a \dashrightarrow r_{a,\mathcal{S}_a}^{R^*}$, and similar to Case (1), node s transmits to nodes $f_{\mathcal{S}}$, $\forall \mathcal{S} \in \{\text{I,II,III,IV}\}, \mathcal{S} \neq \mathcal{S}_a$ with range at most δ . Then we have $f_{\mathcal{S}} \dashrightarrow r_{a,\mathcal{S}}^{R^*}$, $\forall \mathcal{S} \in \{\text{I,II,III,IV}\}, \mathcal{S} \neq \mathcal{S}_a$. We know that $d(f_{\mathcal{S}}, r_{a,\mathcal{S}}^{R^*}) \leq R^*(a)$, $\forall \mathcal{S} \in \{\text{I,II,III,IV}\}, \mathcal{S} \neq \mathcal{S}_a$. Also, we have $d(a, r_{a,\mathcal{S}_a}^{R^*}) \leq R^*(a)$. Therefore:

$$Cost_a^D \leq (4 \lceil \frac{R^*(a)}{\delta} \rceil + 1) \delta^\alpha. \quad (4.17)$$

This results in:

$$\begin{aligned} \mathcal{D}_a &< (4 \lceil \frac{R^*(a)}{\delta} \rceil + 1) \delta^\alpha - ((\lceil \frac{R^*(a)}{\delta} \rceil - 1) \delta)^\alpha \\ \xrightarrow{\alpha \geq 2} \quad \mathcal{D}_a &< 9\delta^\alpha. \end{aligned} \quad (4.18)$$

The maximum difference between the two costs $cost(R^{dist})$ and $cost(R^*)$ happens if there are three nodes with increased range, and the source node covers more than one node. Without considering the effects that the transmission of data by each of them has on the receivers of the other nodes, we can obtain the following rather loose upper bound on $cost(R^D) - cost(R^*)$.

$$\text{cost}(R^D) - \text{cost}(R^*) < \mathcal{K}\delta^\alpha, \quad (4.19)$$

where $\mathcal{K} \leq 36$. \square

Theorem 11. *For a cross network, if the distance between any two nodes on each segment is less than δ , the nodes with increased range in R^* are within a 4δ -neighborhood of the source, or a 6δ -neighborhood of the intersection.*

Proof. For R^* to be optimal, we must have $\mathcal{D}_a \geq 0$, where \mathcal{D}_a is defined in the proof of Theorem 10. For the node a with increased range located on one of the Segments I or II, when it has receivers just on Segments I and II (while transmitting with $R^*(a)$), we have (based on Equation (4.12)):

$$\begin{aligned} & (2\lceil \frac{R^*(a)}{\delta} \rceil + 1)\delta^\alpha - ((\lceil \frac{R^*(a)}{\delta} \rceil - 1)\delta)^\alpha \geq 0 \\ \xrightarrow{\alpha \geq 2} \quad & R^*(a) \leq 4\delta. \end{aligned} \quad (4.20)$$

For node a to have increased range and cover some nodes on segment $\mathcal{S} \in \{\text{I,II}\}$, where $\mathcal{S} \neq \mathcal{S}_a$, we must have $R^*(a) > d(s, a)$. Hence, we must have $d(s, a) < 4\delta$.

For all the other cases, we can use Inequality (4.13), and have:

$$\begin{aligned} & (4\lceil \frac{R^*(a)}{\delta} \rceil + 1 + (\sqrt{2})^\alpha)\delta^\alpha - ((\lceil \frac{R^*(a)}{\delta} \rceil - 1)\delta)^\alpha \geq 0 \\ \xrightarrow{2 \leq \alpha \leq 6} \quad & R^*(a) \leq 6\delta. \end{aligned} \quad (4.21)$$

Node a has some receivers on the other side of the intersection, if $R^*(a) > d(\mathbf{o}, a)$, where \mathbf{o} denotes the position of the intersection. Hence, we must have $d(\mathbf{o}, a) < 6\delta$. \square

4.5.2 Step 2:

Assume a cross network with total length L , in which N nodes are distributed uniformly on all of its segments. The total length of the network is equal to the sum of all segment lengths. For a cross network, we assume that the nodes are first uniformly

distributed on a linear network with length L , and then the total length is subdivided into the number of segments. The distance between any two nodes on each line of the cross network is less than δ , if we do not have an empty δ -length section with that probability. The analysis will use some results of the occupancy theory [71], which are presented next.

The occupancy problem can be described as follows. Assume we have C cells, and N balls to be thrown independently in the cells. The number of empty cells after all the balls have been thrown is a random variable, and we denote it by $\eta(C, N)$. Under the uniform allocation assumption, the following results have been proved [71].

$$P(\eta(C, N) = 0) = \sum_{i=0}^C (-1)^i \binom{C}{i} \left(1 - \frac{i}{C}\right)^N, \quad (4.22)$$

$$E(\eta(C, N)) = C \left(1 - \frac{1}{C}\right)^N, \quad (4.23)$$

$$\text{Var}(\eta(C, N)) = C(C-1) \left(1 - \frac{2}{C}\right)^N + C \left(1 - \frac{1}{C}\right)^N - C^2 \left(1 - \frac{1}{C}\right)^{2N}, \quad (4.24)$$

where $E(\eta(C, N))$ and $\text{Var}(\eta(C, N))$ denote the expected value and the variance of $\eta(C, N)$, respectively.

Suppose we subdivide the whole length of the network, L , into $C = \lceil \frac{L}{\delta/2} \rceil$ cells of length $\delta/2$. With this subdivision, if no empty cells exist, the distance between any two nodes is less than δ .²¹

Using Equation (4.22), we want to have:

$$\sum_{i=0}^C (-1)^i \binom{C}{i} \left(1 - \frac{i}{C}\right)^N \geq 1 - \epsilon. \quad (4.25)$$

For any fixed values of L and $\delta \leq L$, there exist some N values to satisfy Equation (4.25).

To study the asymptotic behavior of $\eta(C, N)$, we focus on the mean and variance of it. According to Equations (4.23) and (4.24), and since

²¹Note that the distance between any two nodes can be less than δ even if some $(\delta/2)$ -length cells are empty, hence the event we consider here is a subset of the event we are interested in. Therefore, the probability of the event we consider is less than what we need, assuring the validity of the results.

$$\delta \leq L \Rightarrow \lceil \frac{2L}{\delta} \rceil \geq 2,$$

the mean and variance of $\eta(C, N)$ converge to zero, when N goes to infinity. This means that $\eta(C, N)$ is a.a.s zero, and δ can be arbitrarily close to zero. Hence, for large number of nodes, the increase in the range of the nodes is not beneficial, and the cost of the distributed assignment (and other assignments that perform better) converges to the cost of the optimal assignment.

4.6 Numerical Results

We study the performance of our proposed algorithms, and the BIP algorithm, by conducting Monte Carlo simulations for different number of nodes on cross networks. We also perform the sweep procedure on the BIP algorithm. But since this procedure has time complexity $\mathcal{O}(N^2)$, to keep the time complexity of the proposed near-optimal and distributed assignments as low as possible, we do not perform the sweep procedure on these assignments. The time complexity of the optimal algorithm and the algorithm to obtain the lower bound on the consumed energy of the optimal algorithm is high, and performing these algorithms on networks with large number of nodes is not practical. We thus present our results for two relatively small values of N , i.e., $N = 14$ and $N = 18$, in the case where the source node is at the intersection, and for $N = 13$ in the general case, where the source node and its location are chosen randomly. For each value of N , we simulate 100 networks. For each network, the nodes are distributed uniformly at random on the cross. The simulation results are summarized in Tables 4.1 and 4.2 for the case where source is at the intersection and the general case, respectively. To make the comparisons easier, we normalize the costs by dividing them by the cost of the optimal assignment in each case. The numbers shown in this table are the average of the normalized costs plus minus the 95% confidence interval. The cost of the BIP algorithm and the BIP algorithm with the sweeping procedure are denoted by $cost(R^{BIP})$ and $cost(R^{BIP/sw})$, respectively. In all simulations, the value of α is selected to be 2.

As Table 4.1 shows, the energy consumption of the near-optimal assignment is close

Table 4.1: The simulation results of cross networks with source at intersection.

	$\frac{\text{cost}(R^{NO})}{\text{cost}(R^*)}$	$\frac{\text{cost}(R^{BIP/sw})}{\text{cost}(R^*)}$	$\frac{\text{cost}(R^{BIP})}{\text{cost}(R^*)}$	$\frac{\text{cost}(R^D)}{\text{cost}(R^*)}$
$N = 14$	1.1140 ± 0.0276	1.2244 ± 0.0364	1.3009 ± 0.0391	1.4303 ± 0.0554
$N = 18$	1.1102 ± 0.0251	1.2100 ± 0.0338	1.2623 ± 0.0360	1.3666 ± 0.0412

Table 4.2: The simulation results of general cross networks.

	$\frac{\text{cost}(R^{low})}{\text{cost}(R^*)}$	$\frac{\text{cost}(R^{NO})}{\text{cost}(R^*)}$	$\frac{\text{cost}(R^{BIP/sw})}{\text{cost}(R^*)}$	$\frac{\text{cost}(R^{BIP})}{\text{cost}(R^*)}$	$\frac{\text{cost}(R^D)}{\text{cost}(R^*)}$
$N = 13$	0.9925 ± 0.0049	1.0668 ± 0.0293	1.1302 ± 0.0362	1.1747 ± 0.0401	1.2556 ± 0.0584

to that of the optimal assignment. Table 4.1 also shows that the proposed near-optimal assignment outperforms BIP and BIP with sweep rather considerably. This is in addition to the advantage of having a lower time complexity. Based on Table 4.1, the distributed assignment performs the worst, but still provides a very low time complexity alternative at the cost of about 40% extra energy compared to the optimal assignment. According to Theorem 10 and the results of Section 4.5.2, one expects the gap to disappear as N tends to infinity. Increasing N from 14 to 18 makes the energy gap between the optimal algorithm and the other sub-optimal algorithms shrink by a non-negligible amount. In particular, one should also note that the performance gap of the distributed algorithm relative to the optimal solution shrinks faster than the other assignments by increasing the size of the network.

For the general case, based on the results of Table 4.2, the difference between the optimal algorithm, on the one hand, and the near-optimal and distributed algorithms, on the other hand, is even less. From Table 4.2 we can see that the calculated lower bound using the addition of one node to the intersection is very tight. Table 4.2 shows that the proposed near-optimal assignment still outperforms both versions of BIP. For larger networks, where the optimal assignment is too complex to find, we only present the results for the other assignments. In Fig. 4.6, we compare the total consumed energy of different assignments normalized with respect to the proposed near-optimal assignment. For each simulation point corresponding to a given number of nodes, 10,000 random networks are generated, each having uniform distribution for the nodes. We run the algorithms on exactly the same networks and obtain the average of the total consumed energy over the 10,000 networks.

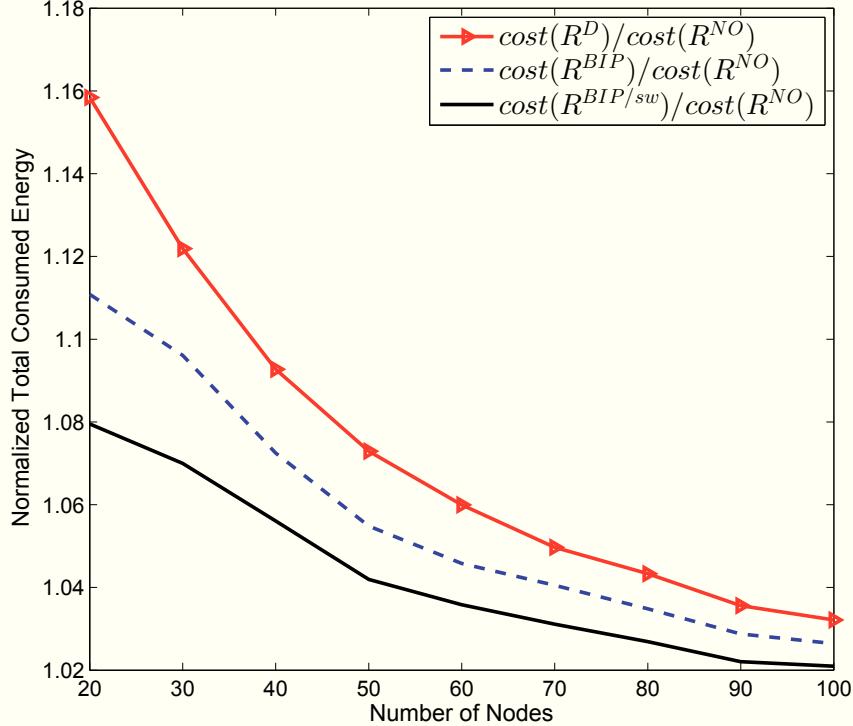


Figure 4.6: Energy consumption comparison of different range assignments.

Fig. 4.6 confirms the superiority of the proposed near-optimal assignment compared with BIP and BIP with sweep. It also shows that the distributed assignment performs close to the other assignments particularly for larger networks. In general, the gap between different algorithms shrinks as the size of the networks increases.

To compare the distributed algorithm with BIP (with and without sweeping procedure) for the general grid networks, we consider a two by two square grid network. We run the algorithms on exactly the same networks and obtain the average of the total consumed energy over 10,000 networks. The comparison between the cost of the distributed algorithm and BIP (with and without sweeping procedure) is shown in Fig. 4.7. It can be seen that for larger number of nodes in the network, the difference between the algorithms is smaller. In particular the difference in energy consumption of the proposed distributed algorithm and BIP with sweep is less than 5% for networks of size $N = 40$ or larger.

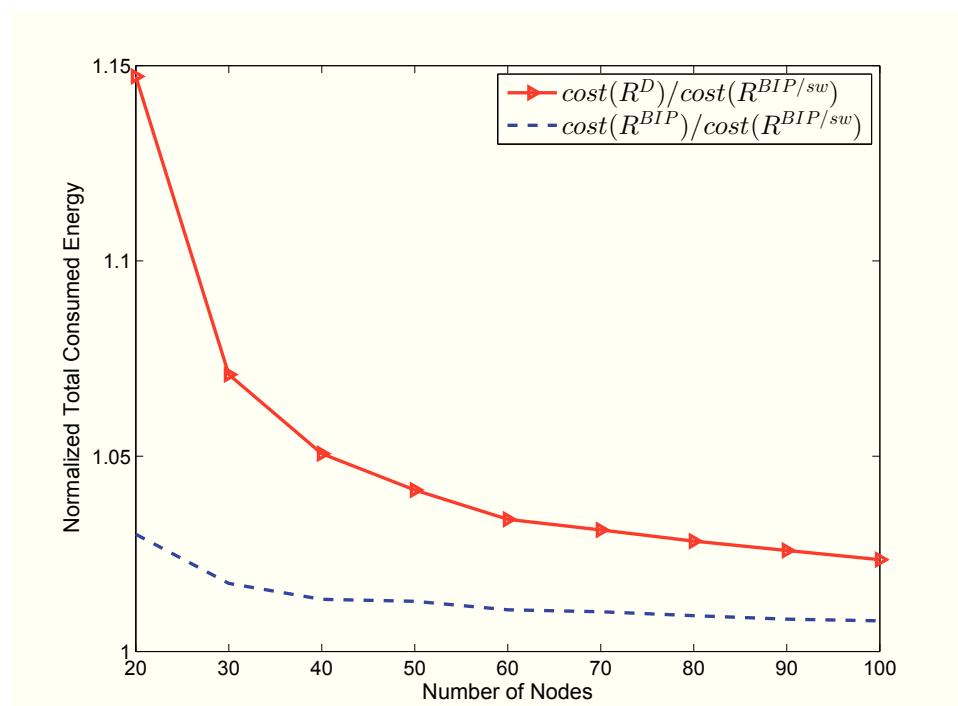


Figure 4.7: Comparison of the energy consumption of different range assignments for two by two square gird networks.

Chapter 5

Probabilistic Linear Networks: Range Assignment

In the remainder of this thesis we consider linear wireless networks, where the nodes are distributed according to a non-homogeneous Poisson distribution with an arbitrary density function. In this chapter, we propose a transmission range assignment for such linear wireless networks. The proposed transmission range assignment is a function of the location on the line and is devised such that there is a high probability of connectivity to the source (located at one end of the network) along the network, while having a low average power consumption for the network¹. This is achieved by maintaining the same (high) probability of having at least one node in the transmission range of the nodes located anywhere in the network. This is equivalent to having the same average number of nodes \mathcal{K} in the transmission range of a node located in an arbitrary location of the network. Our results demonstrate that, for a given average power consumption, the proposed assignment can achieve a higher probability of connectivity compared with an identical transmission range assignment, commonly used in the literature. Furthermore, the number of simultaneous signal receptions in any location of the network, indicating the contention level, is studied in Section 5.4. It is shown that our approach, on average, has a contention level bounded above by \mathcal{K} , while in the identical assignment, this quantity can be larger than \mathcal{K} . The contents of this chapter have been published as [72].

¹We were not able to prove that this assignment consumes the lowest possible energy for a fixed probability of connectivity to the source along the network.

5.1 System Model

We consider a one-dimensional wireless ad hoc network of length L with a fixed transmitter and a fixed receiver located at positions 0 and L , called the source and the destination, respectively. The source transmits the data along the network, and we are interested in connectivity of the network, i.e., the reception of the data by the destination. The total number of nodes in the network and their exact locations are assumed to be unknown. We rather use a probabilistic model and assume a distribution for the network nodes. In this work, similar to [1], we consider a non-homogeneous Poisson process with an arbitrary stationary density function $\lambda(x)$ to describe the distribution of the nodes along the network, where $0 < x < L$ specifies a location in the network (see, e.g., Def. 5.4 of [73]).² This process is referred to as the *network process* in the sequel. We however note that the idea behind the proposed range assignment does not depend on the distribution of the nodes and would be in principle applicable to network processes other than the one considered here, as long as such processes are stationary in time.

We denote by $R(x)$ the transmission range assignment for the position x ($0 \leq x < L$). This means that if a node is located at position x , it will adjust its power based on $R(x)$ to transmit the data. Since the flow of the data is from the source (at location 0) to the destination (at location L), if location x receives the data, all the locations between the source and x have already received the data. Hence, similar to [1], we only consider the coverage of the transmission range of a node towards the destination. So for $y > x$, if $R(x) \geq y - x$, then the node in position y can receive the data from the node in position x ; but if $y - x > R(x)$ and there is no node in the interval $(x, x + R(x)]$, these nodes will not be able to communicate. A position is said to be connected (to the source) if it can receive data from the source (through the network nodes).

Our goal is to find a transmission range assignment $R(x)$ for all $0 \leq x < L$, so that the destination (and hence the whole network) is connected to the source with high probability, while maintaining a low average consumed power for the network.

²Consider a case where the number of nodes in the network is a Poisson random variable with mean $\bar{\lambda}$. Also assume that, given the number of nodes, each node is independently distributed according to a known pdf $f(x)$. Then, in this case, the distribution of the nodes along the network can be modeled by a non-homogeneous Poisson process with the stationary density function $\lambda(x) = f(x)\bar{\lambda}$ [73]. A network realization can thus be obtained, for example, for the purpose of simulations, by first acquiring N , the number of the nodes, according to the Poisson distribution, and then distributing the nodes one by one, according to the pdf $f(x)$ along the network, independently.

In this work, similar to [6, 25, 33–35, 44–48, 54, 55], we do not consider the effects of interference caused by wireless communication among the nodes. It is thus assumed that the network is equipped with a proper multiple access control (MAC) scheduling to avoid interferences. An example of such a MAC scheme is *Minimum Connected Dominating Set* (MCDS)³ broadcasting [27, 74]. In MCDS-broadcasting, only the nodes in the MCDS retransmit the data. There are no polynomial time algorithms to find the MCDS for two dimensional networks [74]. On the other hand for a linear network, the MCDS can be recursively obtained by starting from the source and including in the MCDS, step by step, the farthest node within the transmission range of the previously inserted node.

Nevertheless, we study the number of simultaneous signal receptions in any location of the network, referred to as *contention level*, for the proposed range assignment in comparison with the identical transmission range assignment. In general, lower contention levels are desirable as they result in simpler scheduling schemes and smaller transmission delays.

For simulations, we consider a discrete version of the network process as described in the following. The network length L is partitioned into M segments, each of size $\Delta x = L/M$. We also assume that the source and the destination are located in segments 0 and $M + 1$, respectively. The length Δx is selected to be small enough, i.e., in the order of the average size of the nodes, so that the existence of a node in each segment m , for $1 \leq m \leq M$, is well-approximated by a Bernoulli random variable with the probability of success equal to $P_{OC}[m] = \lambda(x_m) \times \Delta x$, where x_m is the midpoint of segment m , and $P_{OC}[m]$ is the notation for the probability that segment m is occupied by a node. In accordance with the underlying Poisson distribution, such variables are assumed to be independent for different segments [73]. Moreover, we have $P_{OC}[0] = P_{OC}[M + 1] = 1$.

For this discretized network, we use $R_D[m]$ to represent the integer transmission range assigned to segment m . If a node is located in segment m and transmits the data with a transmission range $R_D[m]$, then the data is received in $R_D[m]$ segments towards the destination. Any node located in such segments then receives the data (and is thus named *connected*), and will in turn transmit the data towards the destination at a range which is determined by its location (segment). We say a segment is connected if it is within the transmission range of a connected node. Hence, if there is a node in

³The MCDS is the subset of connected nodes with minimum cardinality such that each node in the network is connected to a node in this subset.

a connected segment, this node can receive the data from the source. In Section 5.3, we calculate the probability of connectivity for all the segments along the network.

5.2 Transmission Range Assignment

For a general (non-uniform) density function $\lambda(x)$, the average distance between adjacent nodes in the network is not a constant and depends on the location x . In fact, for the non-homogeneous Poisson process, the average distance between the adjacent nodes is inversely proportional to $\lambda(x)$. This motivates the search for non-identical transmission range assignments to improve the probability of connectivity for a given average total power consumption in the network. Intuitively, the closer the adjacent nodes, the smaller the transmission range should be. Note that in the identical transmission range assignment, assumed in the existing literature including [1], to maintain a certain probability of connectivity throughout the network, the parts of the network with the lowest density of nodes are the bottlenecks. In such parts, larger transmission ranges are required to maintain the network connectivity. In identical transmission range assignment, this forces a larger transmission range for the whole network, which consequently translates into a waste of power in densely populated parts of the network. Our heuristic approach to have a location-dependent transmission range assignment defines the transmission range assignment for a location x so that at least one node exists in the transmission range of that location (i.e., in the interval $[x, x + R(x)]$) with high probability p_x . The justification is to cover at least one node that can in turn relay the data towards the destination with high probability. The probability that at least one node exists in the interval $[x, x + R(x)]$ is as follows [73]:

$$p_x = 1 - e^{- \int_x^{x+R(x)} \lambda(y) dy}, \quad (5.1)$$

for any location $0 \leq x < L$.

To maintain the same p_x for all locations of the network, we need $\int_x^{x+R(x)} \lambda(y) dy$ to be a constant \mathcal{K} , for all values $0 \leq x < L$. This integral is the average number of nodes in the interval $[x, x + R(x)]$ [73], and thus the proposed range assignment $R(x)$ needs to satisfy:

$$E[N_x] = \int_{x'=x}^{x+R(x)} \lambda(x') dx' = \mathcal{K}, \quad 0 \leq x < L, \quad (5.2)$$

where $E[N_x]$ represents the expected value of N_x , the random variable which denotes the number of nodes that can receive the data from a node at position x with transmission range $R(x)$. Parameter \mathcal{K} dictates the average total consumed power. Since $p_x = 1 - e^{-\mathcal{K}}$, the larger the value of \mathcal{K} , the larger the value of p_x , and hence the higher the probability of connectivity, which comes at the cost of higher power consumption.⁴ Fig. 5.1 illustrates the network model and the proposed transmission range assignment.

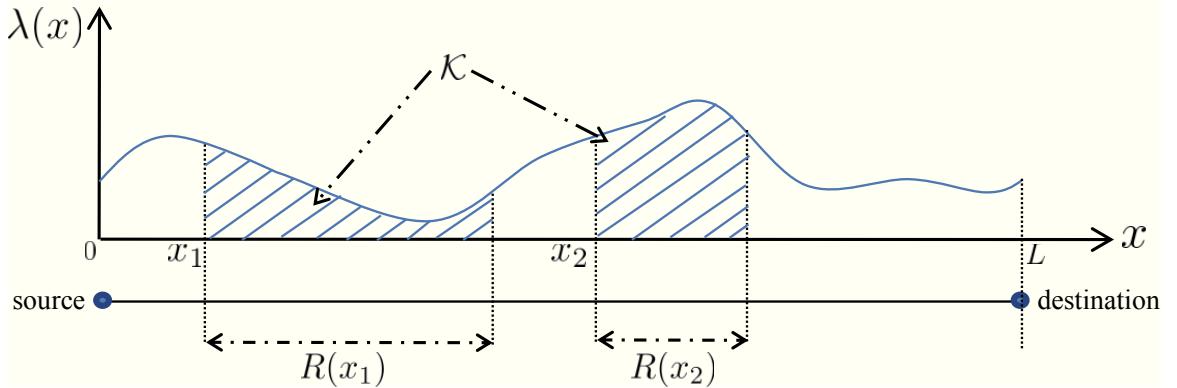


Figure 5.1: Network model and the proposed transmission range assignment.

The range assignment $R(x)$ corresponding to Equation (5.2) is as follows:

$$R(x) = \min\{y_{\mathcal{K}}(x) - x, L - x\}, \quad \text{for } 0 \leq x < L, \quad (5.3)$$

where $L - x$ is used to limit the range so that no power is wasted by communicating beyond the destination, and $y_{\mathcal{K}}(x)$ is the smallest value that satisfies the following equation:

⁴In the selection of \mathcal{K} , one may have to also consider other constraints such as upper and lower bounds on the transmission power of nodes, or the minimum required received power at the nodes. In this work, however, similar to [1], our focus is on the trade-off between the power consumption and network connectivity.

$$\int_{x'=x}^{y_{\mathcal{K}}(x)} \lambda(x') dx' = \mathcal{K}. \quad (5.4)$$

In the discrete version of the network, our proposed assignment needs to satisfy the following equation:

$$E[N_m] = \sum_{j=m+1}^{m+R_D[m]} P_{OC}[j] = \mathcal{K}, \quad (5.5)$$

where the random variable N_m denotes the number of nodes that can receive the data from a node in segment m , while this node's transmission range is $R_D[m]$.

It should be noted that since the variable m in Equation (5.5) is discrete, one may not be able to satisfy all the M equations in Equation (5.5) with exact equality. We therefore choose the smallest integer $R_D[m]$ that satisfies $E[N_m] \geq \mathcal{K}$. The proposed range assignment is thus:

$$R_D[m] = \min \left\{ \arg \min_{m'} \left\{ \sum_{j=m+1}^{m'} P_{OC}[j] \geq \mathcal{K} \right\} - m, M - m \right\}, \text{ for } 0 \leq m \leq M, \quad (5.6)$$

It is easy to see that the time complexity of the proposed transmission range assignment is $\mathcal{O}(M)$, while the uniform transmission range assignment used in [1] has $\mathcal{O}(1)$ time complexity. One should however note that since the range assignment may be performed centrally, its time complexity would not be a burden on the network nodes. In the following, we assume that the expected consumed power of the range assignment $R_D[m]$ can be calculated, up to a constant multiplicative factor, using the following equation:

$$\mathcal{P}(R_D) = \sum_{m=0}^M P_{OC}[m] R_D^\alpha[m], \quad (5.7)$$

where α is the *path-loss exponent* and its value is normally in the range of 2 to 6 [3]. In Equation (5.7), $R_D^\alpha[m]$ represents the power consumption of a node located in

segment m . The average power consumption in segment m is thus represented by $P_{OC}[m]R_D^\alpha[m]$.

5.3 Probability of Connectivity

To have a criteria to compare our approach with others, we calculate the probability of connectivity of the network locations to the source. Here, we present a discrete formula to calculate the probability of connectivity. The more general formula of probability of connectivity, which includes the continuous networks is given in the Chapter 6. Based on an approach similar to the one used in [1], by using the probability that segment m is disconnected, we calculate the probability of connectivity for segment m , $P_D^C[m]$, of the network. Segment m is disconnected if segment $m - 1$ is disconnected, or if segment $m - 1$ is connected but no node transmits beyond segment $m - 1$. Hence, $P_D^C[m]$ can be derived recursively as:⁵

$$P_D^C[m] = P_D^C[m - 1] - \left[1 - \prod_{n \in N_{upto}} (1 - P_{OC}[n]P_D^C[n]) \right] \times \prod_{n' \in N_{over}} (1 - P_{OC}[n']), \quad (5.8)$$

for $1 \leq m \leq M + 1$,

where N_{upto} represents the set of segments n with $R_D[n] + n = m - 1$, and N_{over} denotes the set of segments n' with $R_D[n'] + n' \geq m$. If a node is located in one of the segments of the set N_{upto} , it transmits the data up to segment m (including segment $m - 1$, but not segment m), and if a node occupies one of the segments of the set N_{over} , it transmits beyond segment m . If such a node is connected, segment m will be connected too. For using Equation (5.8) recursively, one needs the initial condition $P_D^C[0] = 1$.

5.4 Contention Analysis

In this section, we analyze the number of concurrent signal receptions at each location x of the network. We denote this parameter by $\mathcal{C}(x)$ and call it the *contention level*

⁵Note that Equation (5.8) for calculating $P_D^C[m]$ is a generalization of the result presented in Lemma 1 of [1] for the special case of $R_D[m] = r, \forall m \in \{0, \dots, M\}$.

at location x , as it indicates the number of nodes that would like their transmitted data to be received at the same time at that location. In the network scenarios considered here, the contention level $\mathcal{C}(x)$ is a non-negative random variable. To clearly define $\mathcal{C}(x)$, we describe the transmission process from the source to the destination in more detail: Consider two sets of nodes \mathbb{T} and \mathbb{R} , with their members being updated sequentially and iteratively as the data passes through the network. The set \mathbb{T} contains the nodes that transmit the data simultaneously, while \mathbb{R} denotes the set of nodes that receive the data (for the first time) simultaneously. At the start of the transmission process, \mathbb{T} just contains the source node, and \mathbb{R} is empty. Once the nodes in \mathbb{T} transmit the data, all the nodes that receive it (simultaneously and for the first time) form the set \mathbb{R} . Then the set \mathbb{T} will be replaced by \mathbb{R} , and \mathbb{R} will be the new set of receivers for \mathbb{T} . Based on the above transmission process, $C(x)$ is defined as the number of transmissions that arrive at location x simultaneously (for the first time). In the following, we show that our proposed transmission range assignment, on average, has a contention level which is bounded above by \mathcal{K} .

Theorem 12. *For the proposed transmission range assignment, the average contention level $E[\mathcal{C}(x)]$ in any location $x, 0 < x \leq L$, of the network is upper bounded by \mathcal{K} .*

Proof of Theorem 12. First, we prove that for any location x in the network, there exists a location $z_x < x$, as the closest location to the source, for which $z_x + R(z_x) \geq x$ and we have $z + R(z) \geq x$ for all values of z in the interval (z_x, x) . The proof is by contradiction. Suppose that there exists a location $z_{x'}$ with the property that $z_{x'} > z_x$, and $z_{x'} + R(z_{x'}) = x' < x$. Using Equations (5.3) and (5.4), we have:

$$\int_{w=z_x}^{z_x+R(z_x)} \lambda(w) dw = \int_{w'=z_{x'}}^{z_{x'}+R(z_{x'})} \lambda(w') dw' = \mathcal{K}. \quad (5.9)$$

Since $z_x < z_{x'}$, and $x' < x$, for Equation (5.9) to hold, $\lambda(w)$ must satisfy $\int_{z_x}^{z_{x'}} \lambda(w) dw = \int_{x'}^x \lambda(w) dw = 0$. As $\lambda(\cdot)$ can be any arbitrary non-negative function, therefore $x = x'$ and $z_x = z_{x'}$, which is in contradiction with the initial assumption. So, for any location x , the nodes whose signals can reach x , are located in $[z_x, x)$. As all the nodes in interval $[z_x, x)$ do not necessarily transmit at the same time, the average number of concurrent signal receptions at any location x satisfies:

$$E[C(x)] \leq \int_{w=z_x}^x \lambda(w) dw = \mathcal{K}. \quad (5.10)$$

□

Using an approach similar to the one used in the proof of Theorem 12, it is easy to see that for the identical transmission range assignment, we have:

$$E[C(x)] \leq \int_{w=x-R_{\text{identical}}}^x \lambda(w) dw, \quad (5.11)$$

where $R_{\text{identical}}$ denotes the value of the identical range. Note that the upper bound in Equation (5.11) can be larger than the upper bound of Equation (5.10) in denser parts of the network. In Section 5.5, we also show that the actual value of $E[C(x)]$ for the identical transmission range assignment can be larger than \mathcal{K} , which itself is an upper bound on $E[C(x)]$ for the proposed assignment.

5.5 Numerical Results

We compare our proposed transmission range assignment with the identical transmission range assignment by comparing the probability of connectivity to the source and the contention level in both cases. For this comparison to be fair, we assume that the average total consumed power for both assignments are equal. For this, first we set the expected number of covered nodes by each node, \mathcal{K} , in our method to be a constant. We choose two values of $\mathcal{K} = 4$ and 6 , which result in having $p_x = 1 - e^{-4} \simeq 0.982$ and $p_x = 1 - e^{-6} \simeq 0.998$, respectively.

Then, using Equation (5.6), we obtain $R_D[m]$ for the proposed scheme. The average total consumed power of the proposed scheme, $\mathcal{P}(R_D)$, is calculated next by Equation (5.7). This is followed by the calculation of the identical transmission range (having the same average total power $\mathcal{P}(R_D)$) as:

$$R_{\text{identical}} = \left[\left(\frac{\mathcal{P}(R_D)}{\sum_{m=0}^{M+1} P_{OC}[m]} \right)^{1/\alpha} \right], \quad (5.12)$$

where $\lceil a \rceil$ is the smallest integer greater than or equal to a .⁶

In our simulations, we discretize the network by selecting $\Delta x = 10$ cm. In the first scenario, we consider $L = 1000$ meters, and we thus have $M = 10,000$ segments. Moreover, we assume that the network nodes are distributed based on the density function $\lambda(x)$ shown in Fig. 5.2. The corresponding range assignment for the proposed scheme is also shown in Fig. 5.2 for two values of $\mathcal{K} = 4$ and 6. The identical transmission range assignment (according to Equation (5.12)) requires the range to be equal to 84.78 and 120.73 meters (corresponding to $R_{\text{identical}} = 8478$ and 12073, respectively) for all the locations, to have the same average total consumed power as the proposed scheme, for the two values of \mathcal{K} , respectively. The probability of connectivity for the two schemes is given in Fig. 5.3, for $\mathcal{K} = 4$ and 6.⁷

As it can be seen, for both the proposed scheme and the identical transmission range assignment, the probability of connectivity improves as \mathcal{K} is increased. This is expected since increasing \mathcal{K} corresponds to larger transmission ranges and higher power consumption. Moreover, for both values of \mathcal{K} , the proposed scheme outperforms the identical transmission range scheme by a large margin over a major part of the network (from about $0.1L$ to L). In particular, the difference between the probability of connectivity of the destination to the source for the two cases is about 35% and 45%, for $\mathcal{K} = 4$ and 6, respectively. Further inspection of Fig. 5.3 also shows that the probability of connectivity to the source drops dramatically after the sparse part of the network, for the identical assignment. The reason is that in this assignment, one has to distribute the power equally to all the locations, without considering the density of the nodes. This results in not having sufficiently large transmission range values for the sparse parts, while having unnecessarily high transmission ranges for the denser parts of the network. For the proposed scheme, the decrease in the probability of connectivity with increasing distance from the source is rather smooth.

⁶The use of $\lceil \cdot \rceil$ in Equation (5.12) results in the total average power to be slightly larger than the target $\mathcal{P}(R_D)$. The increase however is negligible due to the large number of segments in our simulations.

⁷For a fair comparison between the identical and proposed range assignments, we have also reduced the ranges of the positions close to the end of the network for the identical transmission range assignment, so that the coverage of such locations does not go beyond the destination. We have then accordingly increased the range for the rest of the locations to maintain the fixed average total power consumption. As a result, the identical transmission range corresponding to $\mathcal{K} = 4$ and 6 has increased to 87.69 and 126.86 meters, respectively. The results shown in this section are all based on using this method to calculate the identical transmission range.

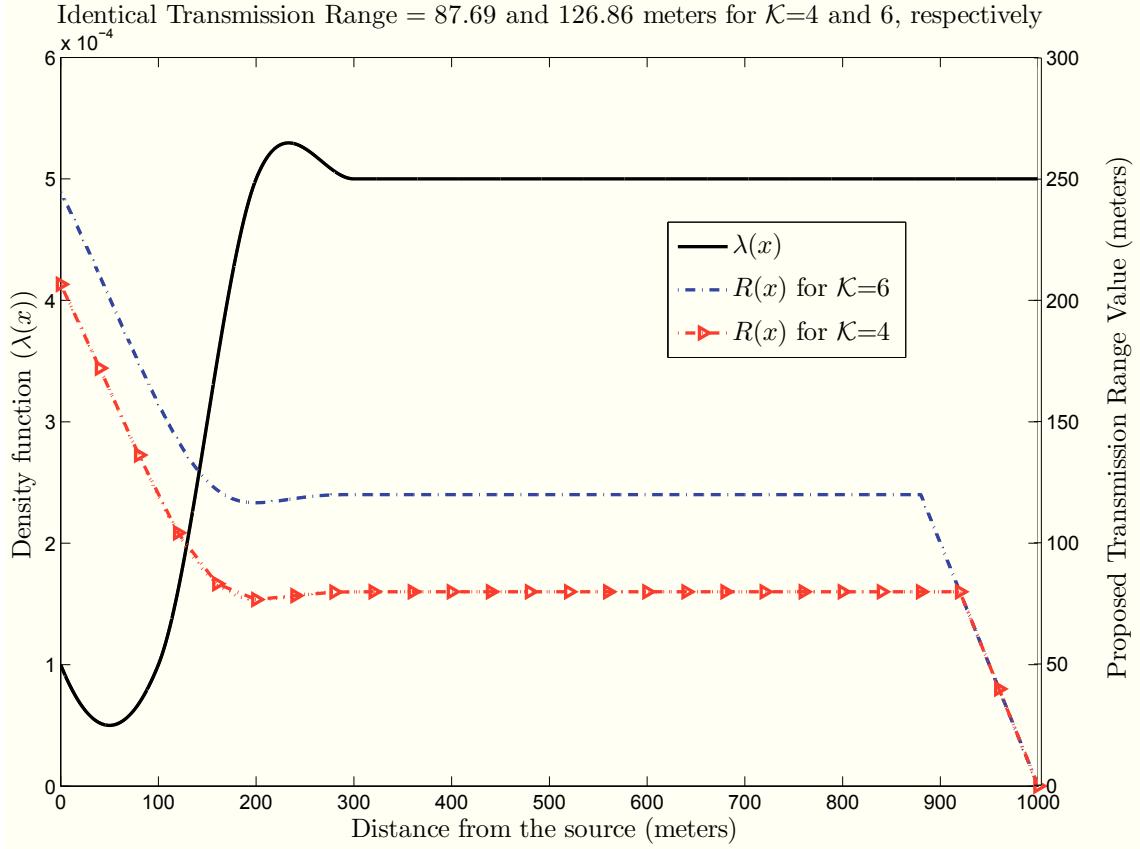


Figure 5.2: Density function $\lambda(x)$ along with the range assignment of the proposed scheme for $K = 4$ and 6 , in the first scenario.

Another important result that can be observed from Fig. 5.3 is the sensitivity of our proposed approach to the density estimation error. Suppose we do not know the density of the network nodes. The simplest estimation is to approximate the density function by a uniform function (having the same average number of nodes). Doing so, we end up having the identical transmission range assignment as the outcome of our proposed approach. This results in (in the worst case) having a much lower probability of connectivity to the source along the network. One should note that although the proposed approach is quite sensitive to the density estimation, it still can perform under limited knowledge, providing a valid transmission range assignment. We also compare the probability of connectivity of the proposed range assignment with that of the identical transmission range assignment for a real-world density function given in Example 4 of [1]. This density function is shown in Fig. 5.4 along

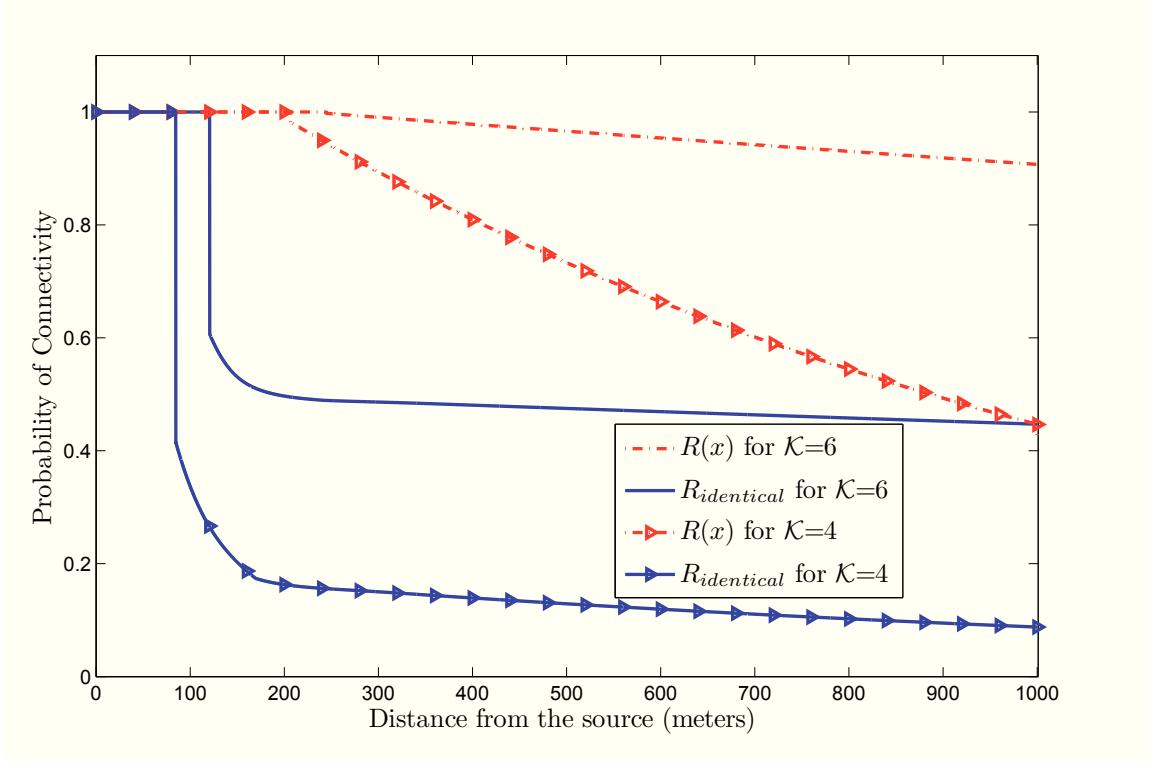


Figure 5.3: Comparison of the probability of connectivity of the network with density function $\lambda(x)$ given in Fig. 5.2 for the two cases of the proposed range assignment $R(x)$ and the identical transmission range assignment R_{idential} for $\mathcal{K} = 4$ and $\mathcal{K} = 6$.

with the range assignment of the proposed scheme for $\mathcal{K} = 6$. In the second scenario, we have $L = 700$ meters and $M = 7,000$ (as $\Delta x = 10$ cm).

The comparison between the probabilities of connectivity of the network for the proposed and the identical transmission range assignments is shown in Fig. 5.5. The value of $P_D^C[m]$ is calculated as the ratio of the number of networks in which segment m is connected to the total number of realized networks (i.e., 10,000). The figure shows that the former consistently outperforms the latter in the range of $x \in [100m, 700m]$ with the maximum difference between the probabilities of connectivity being about 8%.

To compare the contention levels of the proposed range assignment and the identical transmission range assignment, we simulate 10,000 network realizations based on a given density function (using the discrete model of the network and based on the

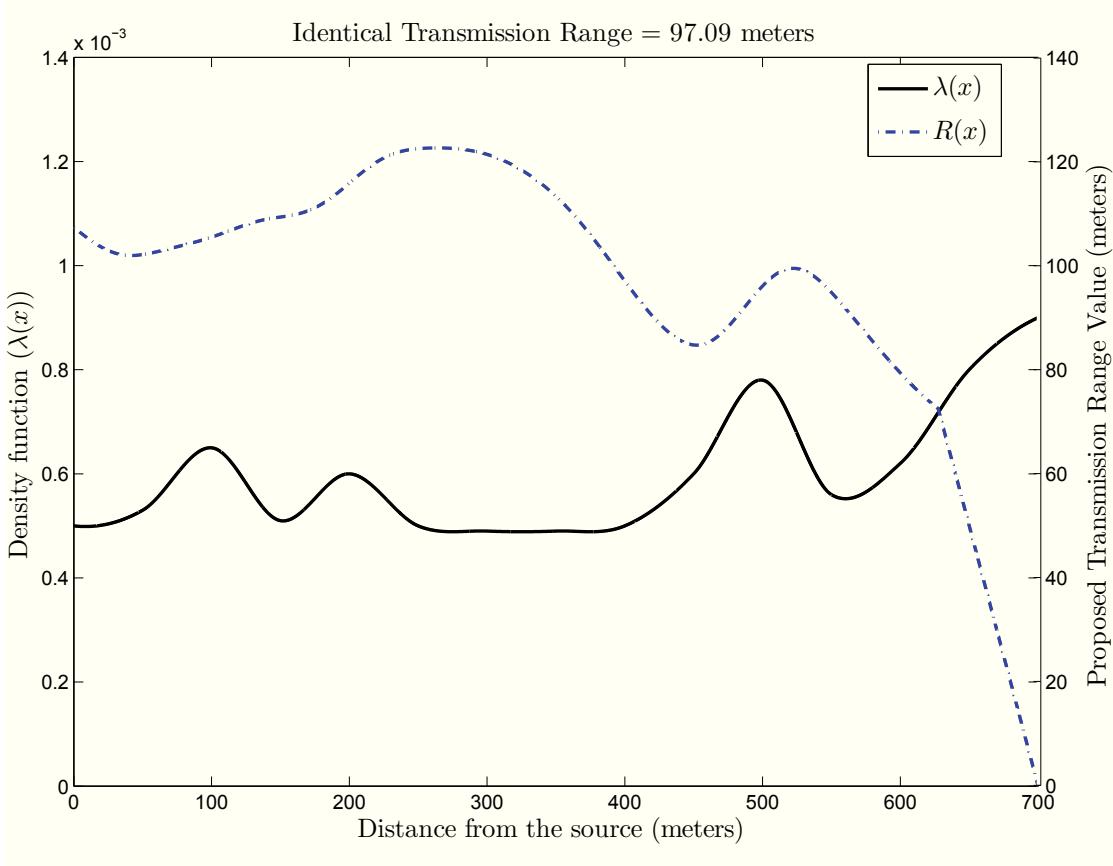


Figure 5.4: Density function given in [1] as a real-world example along with the range assignment of the proposed scheme for $\mathcal{K} = 6$.

technique described in Footnote 2) for each scenario. For each realization, we calculate the contention level $C[m]$ for $0 < m \leq M + 1$, only if the realization is connected (i.e., if the destination can receive the data). For each segment m of the network, we then take the average $\overline{C[m]}$ of $C[m]$ over the connected realizations.

In Fig. 5.6, we have plotted $\overline{C[m]}$ of the proposed range assignment for the density function of Fig. 5.2 and for the two values $\mathcal{K} = 4$ and 6 . In the same figure, we have also reported the corresponding values of $\overline{C[m]}$ for the identical transmission range assignment. As can be seen, for both range assignments, the contention levels generally increase by increasing \mathcal{K} . Note that although the identical transmission range does not depend directly on \mathcal{K} , higher values of \mathcal{K} lead to higher identical transmission ranges, as we want the average total consumed power of both assignments to be equal.

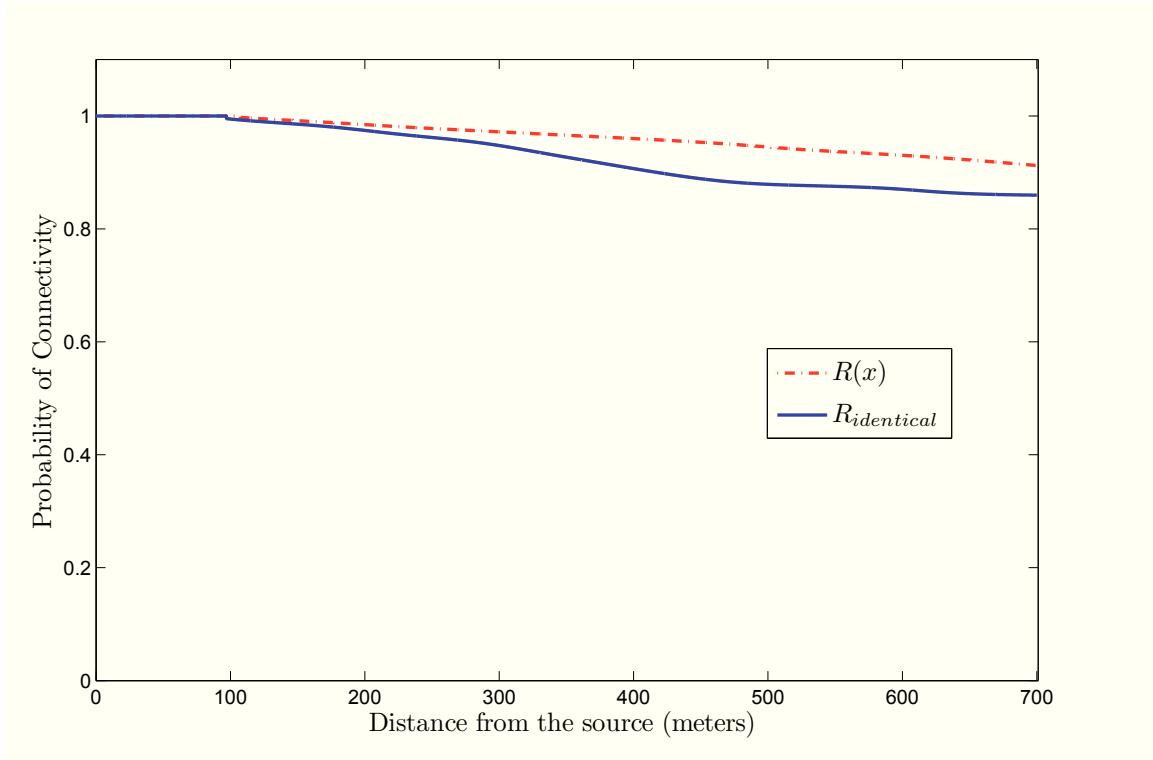


Figure 5.5: Comparison of the probability of connectivity for a real-world example with density function $\lambda(x)$ given in Fig. 5.4 for the two cases of the proposed range assignment $R(x)$ and the identical transmission range assignment $R_{\text{identical}}$ ($\mathcal{K} = 6$).

Fig. 5.6 also shows that although the average contention level for the proposed algorithm is always bounded above by \mathcal{K} , its maximum value is larger than the maximum contention level for the corresponding identical transmission range assignment.

We have also shown $\overline{C[m]}$ of both the proposed and the identical transmission range assignments for the density function of Fig. 5.4 and $\mathcal{K} = 6$ in Fig. 5.7. In this case, the maximum contention value for the identical assignment happens at the destination and is only slightly smaller than the maximum contention for the proposed assignment that occurs at a distance of about 100 meters from the source.

For a case where the contention levels are large for the identical assignment, compared to our scheme, we used a network in which the network nodes are distributed based on the density function $\lambda(x)$ shown in Fig. 5.8. In the third scenario we again have $L = 1000$ meters and $M = 10,000$. The corresponding range assignment for the

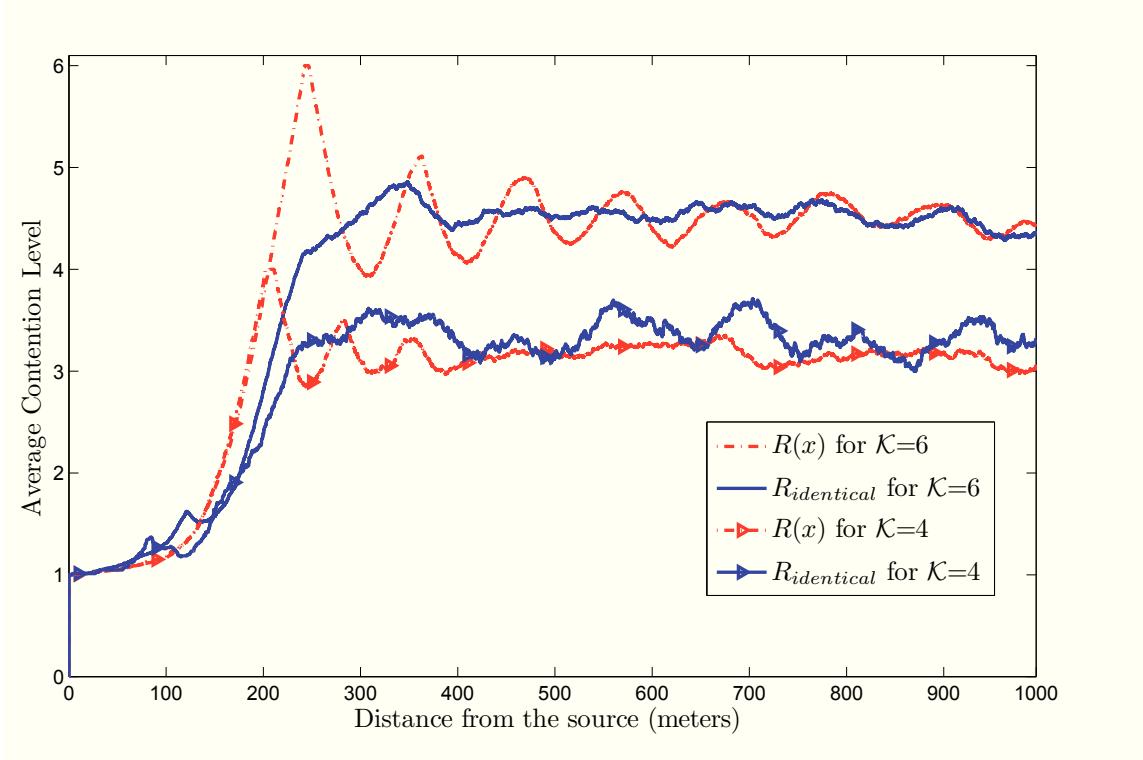


Figure 5.6: Comparison of the average contention levels for the network with the density function $\lambda(x)$ given in Fig. 5.2 for the two cases of the proposed range assignment $R(x)$ and the identical transmission range assignment $R_{\text{identical}}$ ($\mathcal{K} = 4$ and $\mathcal{K} = 6$).

proposed scheme is also shown in Fig. 5.8 for the two values of $\mathcal{K} = 4$ and 6. The identical transmission range assignment given in Equation (5.12) requires the range to be equal to 172.39 and 239.59 meters for all the locations, for the two values of \mathcal{K} , respectively. The probability of connectivity for the two schemes is given in Fig. 5.9, for $\mathcal{K} = 4$ and 6. In Fig. 5.9, Monte Carlo simulations for 10,000 network realizations are also provided. As can be seen, simulations match the analysis very well.

The probability of source-destination connectivity for the proposed scheme however is about 40% and 30% larger than that of identical transmission range assignment for $\mathcal{K} = 4$ and 6, respectively. For the proposed scheme, the decrease in the probability of connectivity with increasing distance from the source is smoother than that of the identical transmission range assignment. We have plotted $\overline{\mathcal{C}[m]}$ for the two assignments for $\mathcal{K} = 4$ and $\mathcal{K} = 6$ of the third scenario in Fig. 5.10. For both cases of

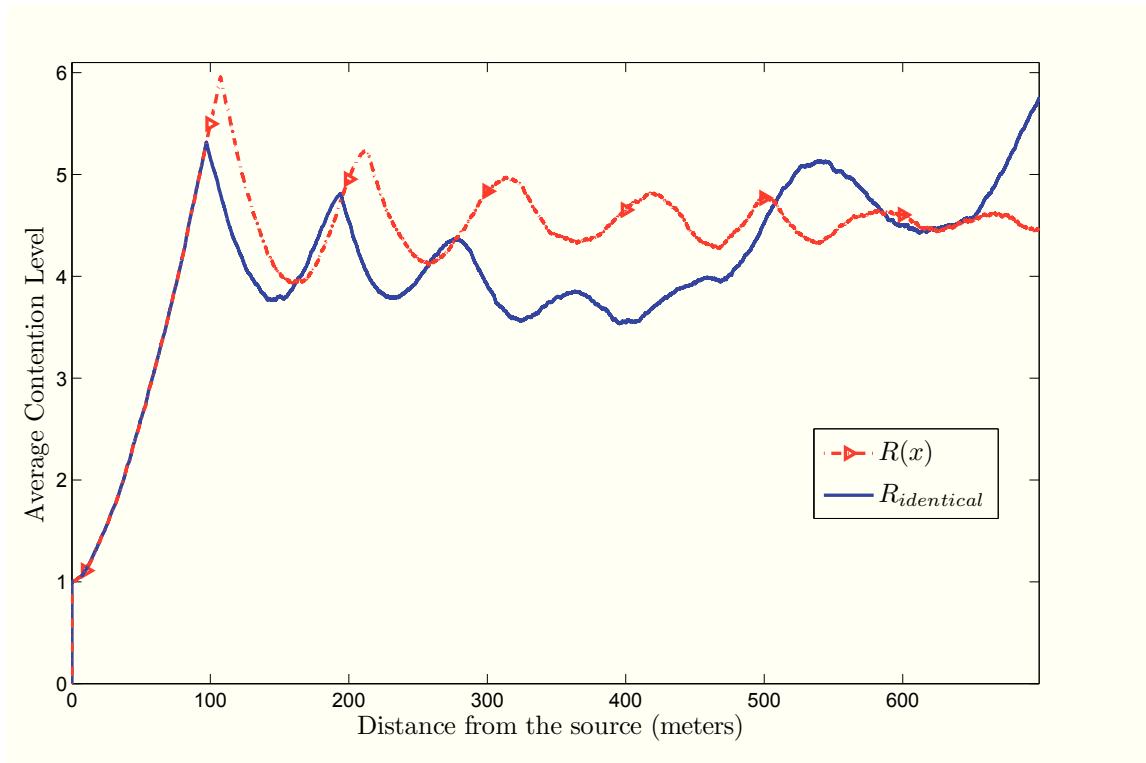


Figure 5.7: Comparison of the average contention levels of the network with the density function $\lambda(x)$ given in Fig. 5.4 for the two cases of the proposed range assignment $R(x)$ and the identical transmission range assignment $R_{\text{identical}}$ ($\mathcal{K} = 6$).

$\mathcal{K} = 4$ and 6 , the maximum average contention level for identical transmission range assignment is more than twice the corresponding values for the proposed scheme.

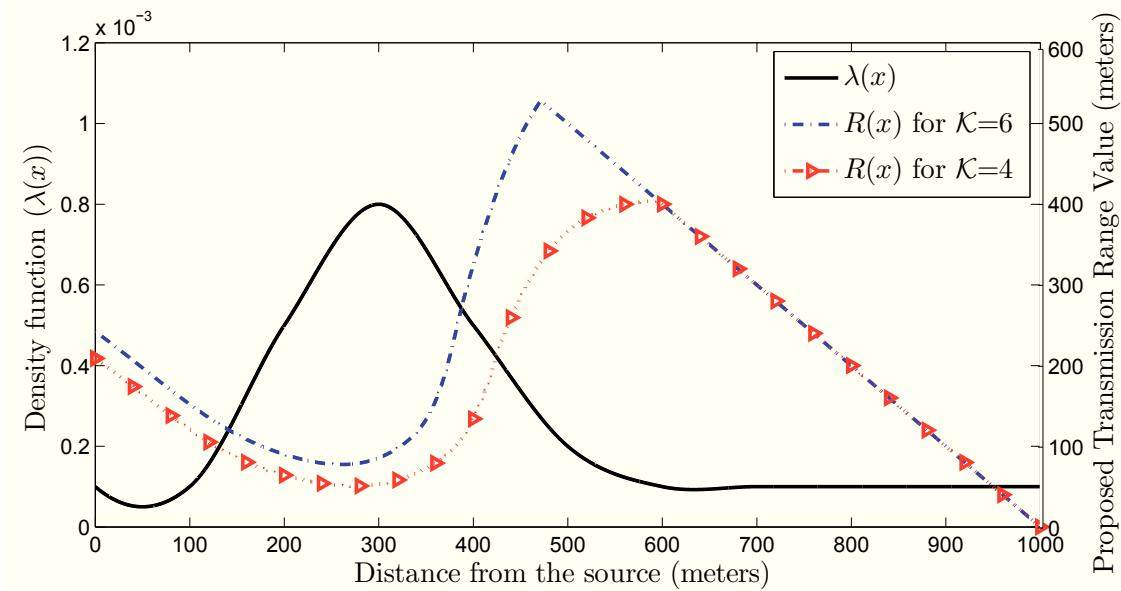


Figure 5.8: Density function $\lambda(x)$ along with the range assignment of the proposed scheme for $\mathcal{K} = 4$ and 6 , in the third scenario.

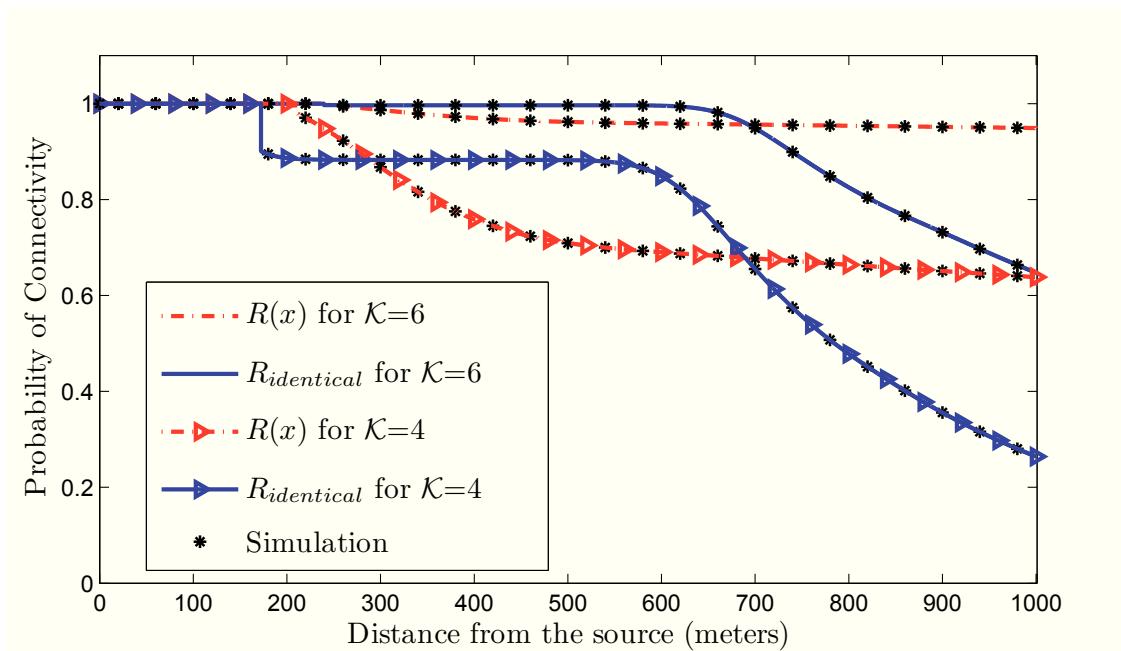


Figure 5.9: Comparison of the probability of connectivity of the network with density function $\lambda(x)$ given in Fig. 5.8 for the two cases of the proposed range assignment $R(x)$ and identical transmission range assignment $R_{\text{identical}}$ for $\mathcal{K} = 4$ and $\mathcal{K} = 6$.

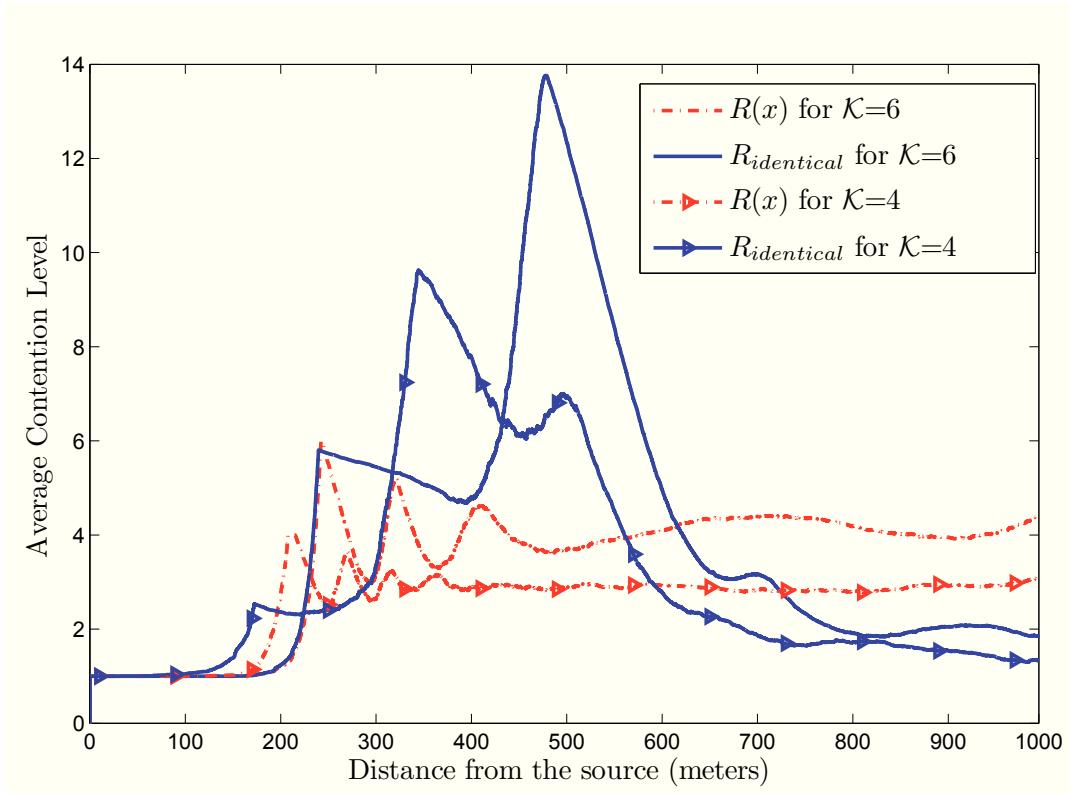


Figure 5.10: Comparison of the average contention levels for the network with density function $\lambda(x)$ given in Fig. 5.8 for the two cases of the proposed range assignment $R(x)$ and identical transmission range assignment $R_{identical}$ for $\mathcal{K} = 4$ and $\mathcal{K} = 6$.

Chapter 6

Probabilistic Linear Networks: Features Analysis

In this chapter, we analyze some features of linear wireless ad hoc networks, where the nodes are placed independently and identically according to Poisson distribution with an arbitrary density function. It is assumed that the source is located at one end of the network, and broadcasts the data to the other end of the network, where the final destination is located. The analyzed features are the probability of connectivity to the source, the number of hops needed for each location of the network to receive the data for the first time (hop-count of that location), and the farthest distance from the source that the data can reach in a given number of hops. The work presented here, to the best of our knowledge, is the first to derive *exact* formulas for the analysis of wireless networks, as described above. Analyzing these features helps to perform delay analysis, which is very important in real-time broadcasting, e.g., broadcasting safety messages, and localization, which can be used in security applications, e.g., location verification. Also, they can be helpful in energy consumption analysis. The main feature of our work is that we consider non-identical transmission range assignments for the nodes. The derived formulas are based on considering a *progressive* transmission range assignment. In a progressive range assignment a range is assigned to every location in the network so that if a node locates there, it transmits the data beyond the assigned transmission range of the locations closer to the source. Finally, we confirm the validity of our analytical approach by simulation.

6.1 System Model

The considered system is exactly the same as what we have in Chapter 5. Moreover, we consider a continuous and a discrete model for the network. The continuous model is more accurate. The discrete model, however, can be useful for simulations, as in most of the literature.

6.1.1 Continuous Network Model

In this model, the probability that an interval I of the network is occupied, i.e., at least one node exists in I , is given by:

$$P_{OC}(I) = 1 - e^{-\int_I \lambda(x) dx}. \quad (6.1)$$

If the interval is very small, we denote it by I_s , and Equation (6.1) can be estimated by:

$$P_{OC}(I_s) \approx \lambda(x_{I_s}) \mathcal{L}(I_s), \quad (6.2)$$

where $\mathcal{L}(I_s)$ denotes the size (or the length) of the interval I_s , and x_{I_s} is the mid-point of I_s . Since the size of I_s is very small, we assume that the density function of the non-homogeneous Poisson process remains a constant in that interval (i.e., $\forall z \in I_s, \lambda(z) = cte$).

In the following, we define a class of transmission ranges that we analyze here.

Definition 10 (Progressive Transmission Range Assignment). *We call a transmission range assignment $R(x), 0 \leq x \leq L$, progressive, if $R(x)$ is continuous in all network locations, and satisfies the following condition: For any network locations x and x' such that $0 \leq x < x' \leq L$, we have $x' + R(x') > x + R(x)$.*

In the networks where the packets are received with no error or loss within the transmission range of the nodes, it is reasonable to assume that an energy-aware transmission range assignment has the property described by Definition 10. Since no data is lost in communication, one-time signal coverage for each position of the network is sufficient. In addition, covering parts that have already been covered is not needed,

as this increases the energy consumption with no benefit. Note that the identical transmission range assignment, used widely in the literature, is progressive. So is the assignment introduced in Chapter 5.

In a network with progressive transmission range assignment $R(x)$, $0 \leq x \leq L$, for any given point $0 \leq x \leq L$, we define location y_x as follows:

$$y_x \triangleq \arg \min_y \{y + R(y) \geq x\} \text{ for } 0 \leq x \leq L. \quad (6.3)$$

Theorem 13. *In a network with progressive transmission range assignment $R(x)$, $0 \leq x \leq L$, for any $x \in [0, L]$, there exists a unique value $y_x \in [0, x]$ such that $y_x = x - R(y_x)$.*

Proof. The existence of y_x can be proved by using the intermediate value theorem [75]¹.

The uniqueness of y_x comes from the fact that since $R(x)$ is progressive, the continuous function $y + R(y)$ is strictly increasing, hence no two y values can have the same $y + R(y)$ value. \square

6.1.2 Discrete Network Model

Similar to [1], we also consider a discrete model of the network. The discrete model asymptotically coincides with the continuous model, as the segment size tends to zero. In the discrete model, we partition the whole network into M segments, each of size $\Delta x = L/M$. We also assume that the source and the destination are located in segments 0 and $M + 1$, respectively. The length Δx is selected to be small enough, i.e., in the order of the average size of the nodes, so that the existence of a node in each segment m , for $1 \leq m \leq M$, is well-approximated by a Bernoulli random variable with the probability of success equal to $P_{OC}[m] = \lambda(x_m)\Delta x$, where x_m is the mid-point of segment m , and $P_{OC}[m]$ is the notation for the probability that segment m is occupied by a node. In accordance with the original Poisson distribution, the variables presenting the occupancy of the segments are assumed to be independent for different segments [73]. Moreover, we have $P_{OC}[0] = P_{OC}[M + 1] = 1$.

¹If a real-valued function $g(\cdot)$ is continuous on the closed interval $[a, b]$ and k is some value between $g(a)$ and $g(b)$, then there exists at least one value c in $[a, b]$ such that $g(c) = k$.

For this discrete network, we use $R_D[m]$ to represent the non-negative integer transmission range assigned to segment m . If a node, which has received the data, is located in segment m and transmits the data with transmission range $R_D[m]$, then the data is received in $R_D[m]$ segments towards the destination. Any node located in such segments then receives the data (and is thus named *connected*), and is able to transmit the data towards the destination at a range which is determined by its location (segment). We say a segment is connected if it is within the transmission range of a connected node. Hence, if there is a node in a connected segment, this node can receive the data from the source.

For the discrete network model, since the transmission ranges are finite integer values rather than real values, the definition of a progressive assignment should be adopted, as follows.

Definition 11 (Discrete Progressive Transmission Range Assignment). *We call a discrete transmission range assignment $R_D[m]$, $0 \leq m \leq M + 1$, progressive, if it satisfies the following condition: For any segments m and m' such that $0 \leq m < m' \leq M + 1$, we have $m + R_D[m] \leq m' + R_D[m']$.*

Unlike the continuous version in which for each location x , only one y_x exists, in the discrete version, for a segment m , we can have multiple (or no) segments n so that $n + R_D[n] = m$. To calculate the desired properties of the network, we introduce the following definitions for progressive transmission range assignments.

Definition 12 (Segment n_m^{\min}). *For any segment $1 \leq m \leq M + 1$, we define the segment n_m^{\min} as the closest segment n to the source such that $n + R_D[n] = m$, i.e.,*

$$n_m^{\min} \triangleq \arg \min_n \{n + R_D[n] = m\}. \quad (6.4)$$

Definition 13 (Segment n_m^{\max}). *For any segment $1 \leq m \leq M + 1$, we define the segment n_m^{\max} as the farthest segment n from the source such that $n + R_D[n] = m$, i.e.,*

$$n_m^{\max} \triangleq \arg \max_n \{n + R_D[n] = m\}. \quad (6.5)$$

Note that there may exist a case where for some segment $1 \leq m \leq M + 1$, there is no segment n that satisfies $n + R_D[n] = m$. In this case, segments n_m^{\min} and n_m^{\max} do not exist. Furthermore, if only one segment n exists so that $n + R_D[n] = m$,

then $n_m^{\min} = n_m^{\max}$. All the segments n that satisfy $n + R_D[n] = m$ for any segment $1 \leq m \leq M + 1$, form the set $\{n_m^{\min}, \dots, n_m^{\max}\}$.

Definition 14 (Segment o_m). *For any segment $1 \leq m \leq M + 1$, we define*

$$o_m \triangleq \arg \min_n \{n + R_D[n] > m\}. \quad (6.6)$$

Definition 15 (Set $\mathcal{N}_m^{\text{cover}}$). *For any segment $1 \leq m \leq M + 1$, we define the set $\mathcal{N}_m^{\text{cover}}$ to contain segments n such that $n < m$ and $n + R_D[n] \geq m$.*

Based on the above definition, all the segments in $\mathcal{N}_m^{\text{cover}}$ can transmit the data to segment m , and if a connected node exists in any of them, segment m will be connected to the source. It is not difficult to see that $\mathcal{N}_m^{\text{cover}} = \{n_m^{\min}, \dots, n_m^{\max}, o_m, \dots, m - 1\}$. We denote the first element of the set $\mathcal{N}_m^{\text{cover}}$ by \mathbf{n}_m . Segment \mathbf{n}_m can be either n_m^{\min} or o_m .

6.2 Probability of Connectivity Analysis

In this section, we derive equations expressing the probability of connectivity of the locations in the network to the source. To the best of our knowledge, the only studies that consider the probability of connectivity in one-dimensional non-homogeneous wireless networks are [1] and our study in Chapter 5. In [1], an identical transmission range assignment was considered and both a discrete and a continuous formula for calculating the probability of connectivity to the source were given. We proposed a formula to calculate the probability of connectivity for any arbitrary discrete transmission range assignment $R_D[m]$, $0 \leq m \leq M + 1$, in Chapter 5.

In this chapter, we present formulas to calculate the probability of connectivity to the source for both the continuous and discrete versions of one-dimensional non-homogeneous wireless networks with progressive transmission range assignments. This generalizes the results of [1] since the identical transmission range assignment is a special case of progressive assignments. Compared to Chapter 5, here we have the continuous and the discrete formulas for deriving the probability of connectivity in the continuous and discrete versions of our network model, and the results presented here are more general.

6.2.1 Probability of Connectivity for the Continuous Network Model

Based on the definition of a progressive transmission range assignment $R(x)$, location x is connected if at least one connected node exists in the interval $[y_x, x)$. To make the calculations simpler, we define the following probability:

$$P^C(x, y) \triangleq P(CN^{[y_x, y]}), \text{ for } y_x \leq y < x, \quad (6.7)$$

where $P(A)$ denotes the probability that event A happens. Notation $CN^{[y_x, y]}$ is used for the event that “at least one connected node exists in interval $[y_x, y]$ ”.

In order to calculate $P^C(x, y)$, we define the variable $f^{[y_x, y]}$ as the location of the first connected node in interval $[y_x, y]$. For this variable, we have the following lemma.

Lemma 17. *For any interval $[v, w]$, $0 \leq v \leq w \leq L$, and any value $z \in [v, w]$, the event $f^{[v, w]} = z$ is equivalent to the event $E^{[v, z)} \cap CN^z$, where $E^{[v, z)}$ denotes the event “the interval $[v, z)$ is empty”, and \cap represents the conjunction of two events. Furthermore, for any two distinct values $z_1, z_2 \in [v, w]$, events $f^{[v, w]} = z_1$ and $f^{[v, w]} = z_2$ are disjoint.*

Proof. First, we need to note the fact that if location $0 \leq z \leq L$ is connected to the source, then all the locations between the source and z are connected to the source as well. If event $f^{[v, w]} = z$ happens, then location z is connected to the source, i.e., event CN^z happens. Also, all the locations in interval $[v, z)$ are connected too. For z to be the first connected node to the source, interval $[v, z)$ has to be empty. That’s because if a node exists in $[v, z)$, it will be the first connected node in interval $[v, w]$. The second claim is obvious as either one of the two distinct values z_1 and z_2 can be the location of the *first* connected node in interval $[v, w]$. \square

The event $CN^{[y_x, y]}$ happens if $f^{[y_x, y]}$ takes a value $z \in [y_x, y]$, as otherwise no connected nodes exist in the interval. Hence $CN^{[y_x, y]}$ is the union of all the events $f^{[y_x, y]} = z$, $\forall z \in [y_x, y]$. According to Lemma 17, these events are disjoint, hence we have:

$$P^C(x, y) = \int_{z=y_x}^y P(f^{[y_x, y]} = z) dz, \text{ for } y_x \leq y < x. \quad (6.8)$$

Based on Lemma 17, we can use the following simplification.

$$\mathrm{P}(f^{[y_x, y]} = z) = \mathrm{P}(E^{[y_x, z]} \cap CN^z) = \mathrm{P}(CN^z | E^{[y_x, z]}) \mathrm{P}(E^{[y_x, z]}). \quad (6.9)$$

The event $CN^z | E^{[y_x, z]}$ means that a connected node exists in location z , while interval $[y_x, z]$ is empty. For this to happen, at least one connected node must exist in interval $[y_z, y_x]$, and a node must exist in location z . This means that $\mathrm{P}(CN^z | E^{[y_x, z]}) = \lambda(z) P^C(z, y_x)$. Therefore, we finally have:

$$\begin{aligned} P^C(x, y) &= \int_{z=y_x}^y \lambda(z) P^C(z, y_x) \mathrm{P}(E^{[y_x, z]}) dz \\ &= \int_{z=y_x}^y \lambda(z) P^C(z, y_x) e^{-\int_{u=y_x}^z \lambda(u) du} dz, \text{ for } y_x \leq y < x. \end{aligned} \quad (6.10)$$

To obtain the probability of connectivity of location x , denoted by $P^C(x)$, we use the following lemma.

Lemma 18. *For any location $0 < x \leq L$, we have $P^C(x) = P^C(x, x)$.*

Proof. Since we assume to have a progressive transmission range assignment, the nodes that can transmit the data to location $0 < x \leq L$ must exist in interval $[y_x, x]$, as there are no nodes before y_x that can cover x . Hence location x is connected if and only if a connected node exists in interval $[y_x, x]$. Thus:

$$\begin{aligned} P^C(x) &= \mathrm{P}(CN^{[y_x, x]}) = P^C(x, x) \\ &= \int_{z=y_x}^x \lambda(z) P^C(z, y_x) e^{-\int_{u=y_x}^z \lambda(u) du} dz, \text{ for } 0 \leq x \leq L. \end{aligned} \quad (6.11)$$

□

To use the continuous formula for finding $P^C(x)$, one needs to solve the integrals numerically. So the continuous formula cannot be used for actual calculations. For that reason, we find the formula in the discrete network model.

6.2.2 Probability of Connectivity for the Discrete Network Model

In the discrete network model with a progressive range assignment, segment m is connected if at least one connected node exists in the set $\mathcal{N}_m^{cover} = \{\mathbf{n}_m, \dots, m-1\}$. Based on the same approach used for calculating the continuous version of the probability of connectivity, we obtain the discrete formula. First, we define:

$$P_D^C[m, n] \triangleq P(CN_D^{\{\mathbf{n}_m, \dots, n\}}), \text{ for } \mathbf{n}_m \leq n < m. \quad (6.12)$$

where $CN_D^{\{\mathbf{n}_m, \dots, n\}}$ denotes the event “at least one connected node exists in the set of segments $\{\mathbf{n}_m, \dots, n\}$ ”.

We define $f_D^{\{\mathbf{n}_m, \dots, n\}}$ as the location of the first connected node in the set of segments $\{\mathbf{n}_m, \dots, n\}$. Similar to the continuous model, we have the following lemma.

Lemma 19. *For any set of segments $\{n_1, \dots, n_T\}$, and any value $k \in \{n_1, \dots, n_T\}$, the event $f_D^{\{n_1, \dots, n_T\}} = k$ is equivalent to the event $E_D^{\{n_1, \dots, k-1\}} \cap CN_D^k$, where $E_D^{\{n_1, \dots, k-1\}}$ denotes the event “the segments in the set $\{n_1, \dots, k-1\}$ are empty”. Furthermore, for any two distinct values $k_1, k_2 \in \{n_1, \dots, n_T\}$, events $f_D^{\{n_1, \dots, n_T\}} = k_1$ and $f_D^{\{n_1, \dots, n_T\}} = k_2$ are disjoint.*

The event $CN_D^{\{\mathbf{n}_m, \dots, n\}}$ is the union of all the events $f_D^{\{\mathbf{n}_m, \dots, n\}} = k \forall k \in \{\mathbf{n}_m, \dots, n\}$.

According to Lemma 19, these events are disjoint, hence we have:

$$\begin{aligned} P_D^C[m, n] &= \sum_{k=\mathbf{n}_m}^n P(f_D^{\{\mathbf{n}_m, \dots, n\}} = k) \\ &= \sum_{k=\mathbf{n}_m}^n P(E_D^{\{\mathbf{n}_m, \dots, k-1\}} \cap CN_D^k) \\ &= \sum_{k=\mathbf{n}_m}^n P(CN_D^k | E_D^{\{\mathbf{n}_m, \dots, k-1\}}) P(E_D^{\{\mathbf{n}_m, \dots, k-1\}}), \text{ for } \mathbf{n}_m \leq n < m. \end{aligned} \quad (6.13)$$

Given that there are no nodes in segments $\{\mathbf{n}_m, \dots, k-1\}$, for segment k to contain a connected node, at least one connected node must exist in the set of segments $\mathcal{N}_k^{cover} \setminus \{\mathbf{n}_m, \dots, k-1\} = \{\mathbf{n}_k, \dots, \mathbf{n}_m-1\}$. Hence $P(CN_D^k | E_D^{\{\mathbf{n}_m, \dots, k-1\}}) = P_{OC}[k] P_D^C[k, \mathbf{n}_m - 1]$. Equation (6.13) can be written as the following:

$$P_D^C[m, n] = \sum_{k=\mathfrak{n}_m}^n \left[P_{OC}[k] P_D^C[k, \mathfrak{n}_m - 1] \prod_{\mathfrak{n}_m \leq j < k} (1 - P_{OC}[j]) \right], \text{ for } \mathfrak{n}_m \leq n < m. \quad (6.14)$$

To obtain the probability of connectivity of segment m to the source, denoted by $P_D^C[m]$, similar to the continuous model, we have the following lemma.

Lemma 20. *For any segment $0 < m \leq M + 1$, we have:*

$$\begin{aligned} P_D^C[m] &= P_D^C[m, m - 1] \\ &= \sum_{k=\mathfrak{n}_m}^{m-1} \left[P_{OC}[k] P_D^C[k, \mathfrak{n}_m - 1] \prod_{\mathfrak{n}_m \leq j < k} (1 - P_{OC}[j]) \right]. \end{aligned} \quad (6.15)$$

According to the above equation, calculation of $P_D^C[m]$ consists of a finite summation over a finite number of functions $P_D^C[m, n]$. By obtaining the values of $P_D^C[m, n]$ starting from the source, we can easily calculate $P_D^C[m]$ for all the segments m in the network. To do that we should consider the below condition:

$$P_D^C[m, n] = 1, \text{ for } 0 < m \leq R_D[0], n < m.$$

6.3 Hop-Count Analysis

In this section, we derive the probability of data reception in a certain number of hops for a specific location of the network. Each hop consists of a group of nodes that receive the data at the same time. We first obtain lower and upper bounds on the distance that the data can reach, given a specific number of hops. Then we derive bounds on the number of hops for the data to reach a specific location. Finally, we present formulas to calculate the probability of the number of hops needed for the data to receive a location (a.k.a, hop-count of that location) for progressive assignments in both the continuous and discrete versions of the network.

6.3.1 Bounds

The bounds we present in this section are valid for any arbitrary transmission range assignment, and not only the progressive ones. Note that the rest of our studies in this chapter are just validated for progressive transmission range assignments. Assume that the network is connected within a given number of hops, h , from the source. Then $d^{\min}[h]$ is the minimum distance from the source that the data can be received for sure. The maximum possible distance for the data to reach in h hops is denoted by $d^{\max}[h]$. These bounds are based on the node distribution and transmission range assignment. We define $d^{\max}[0] = d^{\min}[0] = 0$. As by the first hop, all locations $x \in (0, R(0)]$ receive the data for sure, we have $d^{\max}[1] = d^{\min}[1] = R(0)$. For calculating these bounds, we suppose that the data passes through the network in a progressive way, and does not face a disconnected part in the network (as if it faces a disconnected part, the minimum distance that the data can travel is $R(0)$). Hence $d^{\min}[h] > d^{\min}[h - 1]$ and $d^{\max}[h] > d^{\max}[h - 1]$.

The distance $d^{\max}[h]$ is equal to the maximum distance from the source that a node in $[0, d^{\max}[h - 1]]$ can cover. Hence, $d^{\max}[h]$ is equal to the maximum possible $y + R(y)$ for $0 \leq y \leq d^{\max}[h - 1]$, which is greater than $d^{\max}[h - 1]$. As the maximum of $y + R(y)$ in the interval $[0, d^{\max}[h - 2]]$ is $d^{\max}[h - 1]$, we can reduce the search interval to $(d^{\max}[h - 2], d^{\max}[h - 1])$. Simplifying the limits for y , we come up with the following equation:

$$d^{\max}[h] = \max_{\substack{d^{\max}[h-2] < y \leq d^{\max}[h-1], \\ y + R(y) > d^{\max}[h-1]}} \{y + R(y)\}. \quad (6.16)$$

If $y \leq d^{\max}[h - 2]$, then $y + R(y) \leq d^{\max}[h - 1]$ for sure. Using the inverse of this statement, we have $y + R(y) > d^{\max}[h - 1] \Rightarrow y > d^{\max}[h - 2]$. So by having $y + R(y) > d^{\max}[h - 1]$, the limit $y > d^{\max}[h - 2]$ is already considered. Although the limit $y > d^{\max}[h - 2]$ in the above equation is therefore unnecessary, we still consider it to simplify the search span for y .

To obtain $d^{\min}[h]$, we need to search in all $0 \leq y \leq d^{\max}[h - 1]$ to find $y + R(y)$ greater than $d^{\min}[h - 1]$. But we should note that if $y \in [0, d^{\min}[h - 2]]$, then $y + R(y)$ is already covered in $h - 1$ hops. So we can reduce the search interval to $(d^{\min}[h - 2], d^{\max}[h - 1])$, considering the condition $y + R(y) > d^{\min}[h - 1]$. Using Fig. 6.1, we present an example to show that for general transmission range assignments, the search in $[d^{\min}[h - 2] \leq$

$y \leq d^{\min}[h - 1]$ does not necessarily result in finding $d^{\min}[h]$, and the search span has to encompass up to $d^{\max}[h - 1]$.

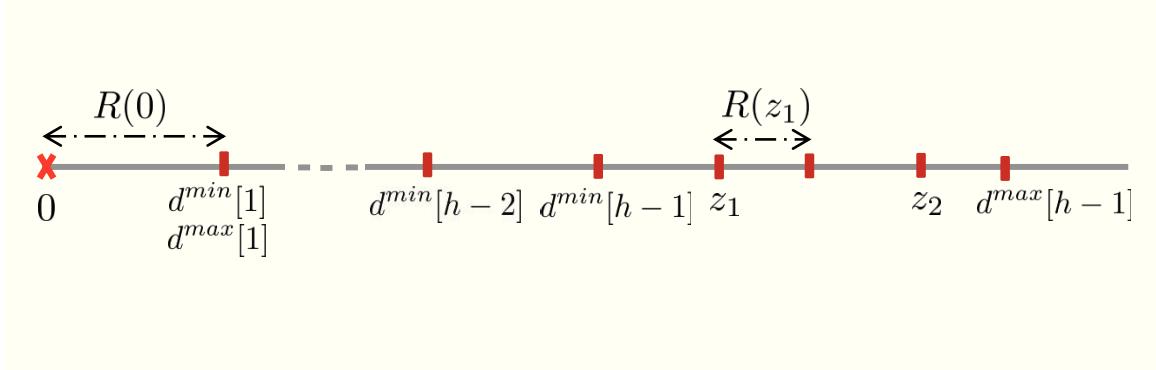


Figure 6.1: A network configuration, to show that the interval $(d^{\min}[h - 2], d^{\max}[h - 1])$ must be searched to find $d^{\min}[h]$.

Suppose for all $y \in (d^{\min}[h - 2], d^{\min}[h - 1])$, we have $y + R(y) \geq z_2$. Since all the locations in interval $[d^{\min}[h - 1], d^{\max}[h - 1]]$ can be reached in $h - 1$ hops, there can exist an $(h - 1)$ -hop path from the source to position z_1 with $z_1 + R(z_1) < z_2$. Hence $d^{\min}[h]$ is not necessarily found by just searching the interval $(d^{\min}[h - 2], d^{\min}[h - 1])$. Thus we have:

$$d^{\min}[h] = \min_{\substack{d^{\min}[h-2] < y \leq d^{\max}[h-1], \\ y + R(y) > d^{\min}[h-1]}} \{y + R(y)\}. \quad (6.17)$$

The following equations give the upper and lower bounds on the number of hops needed for the data to reach a specific location, respectively.

$$\begin{aligned} h^{\max}(x) &= \arg \min_{h'} \{d^{\min}[h'] \geq x\}, \\ h^{\min}(x) &= \arg \min_{h'} \{d^{\max}[h'] \geq x\}. \end{aligned} \quad (6.18)$$

The discrete version of the above bounds are given below:

$$\begin{aligned} h_D^{max}[m] &= \arg \min_{h'} \{d_D^{min}[h'] \geq m\}, \\ h_D^{min}[m] &= \arg \min_{h'} \{d_D^{max}[h'] \geq m\}. \end{aligned} \quad (6.19)$$

where $d_D^{max}[h]$ and $d_D^{min}[h]$ denote the discrete version of $d^{max}[h]$ and $d^{min}[h]$, respectively, which can be derived by:

$$d_D^{max}[h] = \max_{\substack{d_D^{max}[h-2] < n \leq d_D^{max}[h-1], \\ n + R_D[n] > d_D^{max}[h-1]}} \{n + R_D[n]\}. \quad (6.20)$$

$$d_D^{min}[h] = \min_{\substack{d_D^{min}[h-2] < n \leq d_D^{max}[h-1], \\ n + R_D[n] > d_D^{min}[h-1]}} \{n + R_D[n]\}. \quad (6.21)$$

6.3.2 Hop-Count Analysis for the Continuous Network Model

The number of hops required for the data to be received for the first time in location x is called the hop-count of location x and is denoted by $H(x)$. The hop-count for any location of the network is a non-negative integer random variable. As the source node has the data, $H(0) = 0$ (i.e., $P(H(0) = 0) = 1$). The locations in the interval covered by the source's range (i.e., interval $(0, R(0))$) receive the data in exactly one hop, thus we have $P(H(x) = 1) = 1$, for $0 < x \leq R(0)$.

We should emphasize that if $H(x)$ takes some finite integer values with non-zero probabilities, then this location is connected to the source. Thus, we have:

$$\sum_{h=h^{min}(x)}^{h^{max}(x)} P(H(x) = h) = P^C(x), \text{ for } 0 \leq x \leq L. \quad (6.22)$$

To obtain $P(H(x) = h)$ we note that the event $H(x) = h$ happens, if there exists at least one node in interval $[y_x, x)$ with hop-count $h - 1$, and no nodes with hop-count less than $h - 1$ exist in that interval. To simplify the calculations, for any locations

x and y , $y_x \leq y < x \leq L$, we define $\mathbf{H}^h(x, y)$ to be the probability that the event “at least one node exists in interval $[y_x, y]$ with hop-count $h - 1$, and no nodes with hop-count less than $h - 1$ exist in that interval” happens.

Lemma 21. *For any locations x and y , $y_x \leq y < x \leq L$, and any integer number $h > 0$, the probability $\mathbf{H}^h(x, y)$ is equal to the probability of the union of all the events $(f^{[y_x, y]} = z) \cap (H(z) = h - 1)$, $\forall z \in [y_x, y]$.*

Proof. We use the fact that if location $0 \leq z \leq L$ has hop-count $h - 1$, the hop-count for all the locations between z and the destination is greater than or equal to $h - 1$. Assume for a value $z \in [y_x, y]$, event $(f^{[y_x, y]} = z) \cap (H(z) = h - 1)$ happens. Since z is the location of the first connected node with hop-count $h - 1$ in interval $[y_x, y]$, then according to the above fact, no nodes with hop-count less than $h - 1$ exist in that interval. Hence event “at least one node exists in interval $[y_x, y]$ with hop-count $h - 1$, and no nodes with hop-count less than $h - 1$ exist in that interval” happens. But this is true for any value $z \in [y_x, y]$. So to calculate $\mathbf{H}^h(x, y)$, one needs to consider the probability of the union of all the events $(f^{[y_x, y]} = z) \cap (H(z) = h - 1)$, $\forall z \in [y_x, y]$. \square

According to Lemma 17, for any two distinct values $z_1, z_2 \in [y_x, y]$, events $f^{[y_x, y]} = z_1$ and $f^{[y_x, y]} = z_2$ are disjoint. Thus, for any integer number $h > 0$, events $(f^{[y_x, y]} = z_1) \cap (H(z_1) = h - 1)$ and $(f^{[y_x, y]} = z_2) \cap (H(z_2) = h - 1)$ are disjoint too. Using Lemmas 17 and 21, we have:

$$\begin{aligned}\mathbf{H}^h(x, y) &= \int_{z=y_x}^y P((f^{[y_x, y]} = z) \cap (H(z) = h - 1)) dz \\ &= \int_{z=y_x}^y P(E^{[y_x, z]} \cap CN^z \cap (H(z) = h - 1)) dz \\ &= \int_{z=y_x}^y P(CN^z \cap (H(z) = h - 1) | E^{[y_x, z]}) P(E^{[y_x, z]}) dz.\end{aligned}\tag{6.23}$$

Given that the interval $[y_x, z]$ is empty, a node with hop-count $h - 1$ exists in location z if at least one node with hop-count $h - 2$ exists in interval $[y_z, y_x]$, and no nodes with hop-count less than $h - 2$ exist in that interval, meaning $P(CN^z \cap (H(z) = h - 1) | E^{[y_x, z]}) = \lambda(z) \mathbf{H}^{h-1}(z, y_x)$. So:

$$\mathbf{H}^h(x, y) = \int_{z=y_x}^y \lambda(z) \mathbf{H}^{h-1}(z, y_x) e^{-\int_{u=y_x}^z \lambda(u) du} dz, \text{ for } y_x \leq y < x. \quad (6.24)$$

To derive $P(H(x) = h)$, we use the following lemma, which is easy to prove.

Lemma 22. *For any location $0 < x \leq L$, we have $P(H(x) = h) = \mathbf{H}^h(x, x)$.*

Thus we have:

$$P(H(x) = h) = \int_{z=y_x}^x \lambda(z) \mathbf{H}^{h-1}(z, y_x) e^{-\int_{u=y_x}^z \lambda(u) du} dz, \text{ for } 0 < x \leq L. \quad (6.25)$$

To obtain $P(H(x) = h)$, one needs to solve the above integral numerically. So the continuous formula cannot be used for actual calculations. To make calculations possible, we find the formula in the discrete network model.

6.3.3 Hop-Count Analysis for the Discrete Network Model

In this section, we consider the discrete network model with a progressive transmission range assignment. The number of hops required for segment m to receive the data (for the first time) is denoted by $H_D[m]$. Similar to the continuous case, we have $P(H_D[0] = 0) = 1$, and $P(H_D[m] = 1) = 1$, for $1 \leq m \leq R_D[0]$. Also:

$$\sum_{h=h_D^{\min}[m]}^{h_D^{\max}[m]} P(H_D[m] = h) = P_D^C[m], \text{ for } m = 0, \dots, M + 1. \quad (6.26)$$

We have $H_D[m] = h$, if at least one node with hop-count $h - 1$ exists in segments $\mathcal{N}_m^{\text{cover}}$, and no nodes with hop-count less than $h - 1$ exist there. Similar to the approach used for the continuous formula, for any segments m and n , $\mathfrak{n}_m \leq n < m \leq M + 1$, we define $\mathbf{H}_D^h[m, n]$ as the probability that the event “at least one node exists in set of segments $\{\mathfrak{n}_m, \dots, n\}$ with hop-count $h - 1$, and no nodes with hop-count less than $h - 1$ exist in that set of segments”. The following lemma can be proved similar to Lemma 21.

Lemma 23. For any segments m and n , $\mathbf{n}_m \leq n < m \leq M + 1$, and any integer number $h > 0$, the probability $\mathbf{H}_D^h[m, n]$ is equal to the probability of the union of all the events $(f_D^{\{\mathbf{n}_m, \dots, n\}} = k) \cap (H_D[k] = h - 1)$, $\forall k \in \{\mathbf{n}_m, \dots, n\}$.

Since events $f_D^{\{\mathbf{n}_m, \dots, n\}} = k_1$ and $f_D^{\{\mathbf{n}_m, \dots, n\}} = k_2$, for any two distinct segments k_1 and k_2 , are disjoint (according to Lemma 19), events $(f_D^{\{\mathbf{n}_m, \dots, n\}} = k_1) \cap (H_D[k_1] = h - 1)$ and $(f_D^{\{\mathbf{n}_m, \dots, n\}} = k_2) \cap (H_D[k_2] = h - 1)$ for any integer $h > 0$ are disjoint too. Using Lemmas 19 and 23, we have:

$$\begin{aligned} \mathbf{H}_D^h[m, n] &= \sum_{k=\mathbf{n}_m}^n P((f_D^{\{\mathbf{n}_m, \dots, n\}} = k) \cap (H_D[k] = h - 1)) \\ &= \sum_{k=\mathbf{n}_m}^n P(E_D^{\{\mathbf{n}_m, \dots, k-1\}} \cap CN_D^k \cap (H_D[k] = h - 1)) \\ &= \sum_{k=\mathbf{n}_m}^n P(CN_D^k \cap (H_D[k] = h - 1) | E_D^{\{\mathbf{n}_m, \dots, k-1\}}) P(E_D^{\{\mathbf{n}_m, \dots, k-1\}}), \text{ for } \mathbf{n}_m \leq n < m. \end{aligned} \quad (6.27)$$

Given that there are no nodes in segments $\{\mathbf{n}_m, \dots, k-1\}$, for segment k to contain the first node with hop-count $h - 1$, at least one node with hop-count $h - 2$ must exist in the set of segments $\mathcal{N}_k^{cover} \setminus \{\mathbf{n}_m, \dots, k-1\} = \{\mathbf{n}_k, \dots, \mathbf{n}_{m-1}\}$. Thus, $P(CN_D^k \cap (H_D[k] = h - 1) | E_D^{\{\mathbf{n}_m, \dots, k-1\}}) = P_{OC}[k] \mathbf{H}_D^{h-1}[k, \mathbf{n}_{m-1}]$, which results in:

$$\mathbf{H}_D^h[m, n] = \sum_{k=\mathbf{n}_m}^n \left[P_{OC}[k] \mathbf{H}_D^{h-1}[k, \mathbf{n}_{m-1}] \prod_{\mathbf{n}_m \leq j < k} (1 - P_{OC}[j]) \right], \text{ for } \mathbf{n}_m \leq n < m. \quad (6.28)$$

Similar to the continuous model, to derive $P(H_D[m] = h)$, we have the following lemma.

Lemma 24. For any segment $0 \leq m \leq M + 1$, we have:

$$\begin{aligned} P(H_D[m] = h) &= \mathbf{H}_D^h[m, m-1] \\ &= \sum_{k=\mathbf{n}_m}^{m-1} \left[P_{OC}[k] \mathbf{H}_D^{h-1}[k, \mathbf{n}_{m-1}] \prod_{\mathbf{n}_m \leq j < k} (1 - P_{OC}[j]) \right]. \end{aligned} \quad (6.29)$$

The above equation consists of a finite summation over a finite number of functions $\mathbf{H}_D^h[m, n]$. By obtaining the values of $\mathbf{H}_D^h[m, n]$ starting from the source, we can easily calculate $P(H_D[m] = h)$ for all the segments m in the network. To do that we should consider the below condition:

$$\mathbf{H}_D^h[m, n] = 1, \text{ for } 0 < m \leq R_D[0], n < m, h = 1. \quad (6.30)$$

6.4 Last Reachable Distance Analysis

We analyze the last location of the network that receives the data in a given number of hops. Similar to the previous sections, we first analyze the continuous version, and then the discrete version of the network.

6.4.1 Last Reachable Distance Analysis for the Continuous Network Model

In the continuous network model, the last location that receives the data by h hops is denoted by $D[h]$. Since $D[h]$ is a continuous random variable, we need to describe it by its pdf, denoted by $\mathfrak{d}^h(x)$ for all $0 \leq x \leq L$ and $h^{\min}(x) \leq h \leq h^{\max}(x)$. We have $\mathfrak{d}^0(x) = \delta(x)$ and $\mathfrak{d}^1(x) = \delta(x - R(0))$, where $\delta(x)$ is the Dirac delta function. For a network with progressive transmission range assignment, location x is the last reachable location by h hops, if a node exists in location y_x , which has received the data by $h - 1$ hops, and no other nodes exist in interval (y_x, x) that receive the data by $h - 1$ hops. The node located in y_x is the last *node* (*and not location*) that receives the data by $h - 1$ hops, if $D[h - 1]$ is greater than y_x , and $D[h - 2]$ is smaller than y_x , and the interval $(y_x, D[h - 1])$ is empty. To be able to obtain the probability of this event, we use the joint pdf $\mathfrak{d}^{h,h-1}(x, y)$, which denotes the joint probability density function of random variables $D[h]$ at point x , and $D[h - 1]$ at point $y_x \leq y < x$.

$$\mathfrak{d}^{h,h-1}(x, y) = \lambda(y_x) e^{-\int_{u=y_x}^y \lambda(u) du} \int_{z=d^{\min}[h-1]}^{y_x} \mathfrak{d}^{h-1,h-2}(y, z) dz, \text{ for } y_x \leq y < x. \quad (6.31)$$

Using the marginal probability, the pdf of $D[h]$ is:

$$\begin{aligned}\mathfrak{d}^h(x) &= \int_{y=y_x}^x \mathfrak{d}^{h,h-1}(x,y) dy \\ &= \int_{y=y_x}^x \left(\lambda(y_x) e^{-\int_{u=y_x}^y \lambda(u) du} \int_{z=d^{\min}[h-1]}^{y_x} \mathfrak{d}^{h-1,h-2}(y,z) dz \right) dy.\end{aligned}\quad (6.32)$$

Therefore, for any give number of hops, h , the random variable $D[h]$ is in interval $[d^{\min}[h], d^{\max}[h]]$ with the above derived pdf.

Another important parameter that can be derived using the pdf of the random variable $D[h]$ is the probability of the event that the data has passed h hops from the source towards the destination. We denote this event by \exists^h . The following relation should be noted:

$$P(D[h] = x) = P(\exists^h \cap (D[h] = x)).$$

By calculating the marginal probability, we derive:

$$P(\exists^h) = \int_{x=d^{\min}[h]}^{d^{\max}[h]} \mathfrak{d}^h(x), \text{ for } 0 \leq h \leq h^{\max}(L). \quad (6.33)$$

Using the cumulative density function (cdf) of the random variable $D[h]$, denoted by $\mathfrak{D}^h(x) = \int_{y=0}^x \mathfrak{d}^h(y) dy$, we can rewrite the above equation as:

$$P(\exists^h) = \mathfrak{D}^h(L), \text{ for } 0 \leq h \leq h^{\max}(L), \quad (6.34)$$

considering the fact that $\mathfrak{d}^h(y) = 0$ for $y \notin [d^{\min}[h], d^{\max}[h]]$.

This probability can be used to obtain the last location of the network that receives the data from the source within a specific number of hops. This parameter can be used in data dissemination analysis in VANETs where the portion of the network that receive the data up to a certain number of hops is important. One example are safety applications, where the number of hops can be used as an estimation of the delay needed for a message to propagate [7, 8]. To analyze the distance from the source that the data can disseminate in a limited hop-count (which can be translated into limited time) the below probability can be used.

$$P(D[h] \geq x | \exists^h) = 1 - \frac{P(D[h] \leq x)}{P(\exists^h)} = 1 - \frac{\mathfrak{D}^h(x)}{\mathfrak{D}^h(L)}. \quad (6.35)$$

The above equation gives the probability of the event that the last reachable location by at most h hops² (a fixed maximum delay) is greater than or equal to x .

Similar to the other continuous formulas we found, the calculation of the probabilities concerning the $D[h]$ random variable needs numerical integrations. In the next section we present the discrete formulas for last reachable distance analysis.

6.4.2 Last Reachable Distance Analysis for the Discrete Network Model

In the discrete network model, the last segment that receives the data by h hops is denoted by $D_D[h]$, and is a discrete random variable. As the source has the data, we have $D_D[0] = 0$ (i.e., $P(D_D[0] = 0) = 1$). Also, the last segment to receive the data by 1 hop is $R_D[0]$, so $D_D[1] = R_D[0]$ (i.e., $P(D_D[1] = R_D[0]) = 1$). In networks with progressive transmission range assignments, segment m is the last segment that receives the data by h hops, if it is the last segment in the transmission range of a node which has received the data by $h - 1$ hops, and all the other segments that receive the data with $h - 1$ hops, and are beyond that segment are empty. So, for segment m to be the last covered segment by h hops, at least one node must exist in the set of segments $\{n_m^{min}, \dots, n_m^{max}\}$ which receives the data by $h - 1$ hops, and all the segments in set $\{o_m, \dots, m - 1\}$ that receive the data by $h - 1$ hops must be empty. To obtain this probability, similar to the continuous case, we define the below joint probability:

$$\mathbf{D}^{h,h-1}[m, n] \triangleq P(D[h] = m, D[h - 1] = n) \text{ for } 0 \leq n < m \leq M + 1. \quad (6.36)$$

To calculate $\mathbf{D}^{h,h-1}[m, n]$, we consider the following two cases:

- a) $n_m^{min} \leq n \leq n_m^{max}$. For this case, the event $D[h] = m$ happens, while $D[h - 1] = n$, if a node with hop-count $h - 1$ exists in set of segments $\{n_m^{min}, \dots, n\}$. Note that we can have $D[h - 1] = n$, while $D[h - 2]$ is within the set of segments $\{n_m^{min}, \dots, n - 1\}$.

²Note the term *at most h hops*, as we do not condition x to be greater than $D[h - 1]$.

The below equation can be used to calculate $\mathbf{D}^{h,h-1}[m, n]$ for this case:

$$\begin{aligned}\mathbf{D}^{h,h-1}[m, n] &= \sum_{d_{h-2}=d_D^{min}[h-2]}^{n_m^{min}-1} \left[\mathbf{D}^{h-1,h-2}[n, d_{h-2}] \left(1 - \prod_{n'=n_m^{min}}^n (1 - P_{OC}[n']) \right) \right] \\ &\quad + \sum_{d_{h-2}=n_m^{min}}^{n-1} \left[\mathbf{D}^{h-1,h-2}[n, d_{h-2}] \left(\prod_{n'=n_m^{min}}^{d_{h-2}} (1 - P_{OC}[n']) \right) \left(1 - \prod_{n''=d_{h-2}+1}^n (1 - P_{OC}[n'']) \right) \right].\end{aligned}\tag{6.37}$$

In the second term of the summation in the above equation, we used the fact that no nodes exist in the set of segments $\{n_m^{min}, \dots, d_{h-2}\}$, as if some exist $D[h-1] = m \neq n$, which is unacceptable.

b) $o_m \leq n \leq m-1$: Here, for $D[h] = m$ and $D[h-1] = n$ to happen, in addition to the conditions needed for the above case, there must be no nodes in the set of segments $\{o_m, \dots, n\}$, as if one exists, we have $D[h] > m$. So:

$$\begin{aligned}\mathbf{D}^{h,h-1}[m, n] &= \left\{ \prod_{o=o_m}^n (1 - P_{OC}[o]) \right\} \times \\ &\quad \left\{ \sum_{d_{h-2}=d_D^{min}[h-2]}^{n_m^{min}-1} \left[\mathbf{D}^{h-1,h-2}[n, d_{h-2}] \left(1 - \prod_{n'=n_m^{min}}^{n_m^{max}} (1 - P_{OC}[n']) \right) \right] \right. \\ &\quad \left. + \sum_{d_{h-2}=n_m^{min}}^{n_m^{max}-1} \left[\mathbf{D}^{h-1,h-2}[n, d_{h-2}] \left(\prod_{n'=n_m^{min}}^{d_{h-2}} (1 - P_{OC}[n']) \right) \left(1 - \prod_{n''=d_{h-2}+1}^{n_m^{max}} (1 - P_{OC}[n'']) \right) \right] \right\}.\end{aligned}\tag{6.38}$$

The random variable $D[h]$ is equal to m with the following probability:

$$P(D_D[h] = m) = \begin{cases} \sum_{n=n_m^{min}}^{m-1} \mathbf{D}^{h,h-1}[m, n], & d_D^{min}[h] \leq m \leq d_D^{max}[h], \\ 0, & \text{otherwise.} \end{cases}\tag{6.39}$$

Similar to the continuous version, we have:

$$P(\exists_D^h) = \sum_{n=0}^{M+1} P(D_D[h] = n), \text{ for } 0 \leq h \leq h_D^{max}[M+1], \quad (6.40)$$

in which \exists_D^h denotes the event that the data has passed h hops from the source towards the destination.

For data dissemination analysis in the discrete networks, we have:

$$P(D_D[h] \geq m | \exists_D^h) = 1 - \frac{P(D[h] \leq m)}{P(\exists_D^h)} = 1 - \frac{\sum_{n=0}^m P(D_D[h] = n)}{\sum_{n=0}^{M+1} P(D_D[h] = n)}, \quad (6.41)$$

which gives the probability of the event that the last reachable location by at most h hops is greater than or equal to segment m .

6.4.3 Obtaining the Hop-Count pmf using the Last Reachable Distance Information

Location x receives the data by h hops, if it is within the interval $(D[h-1], D[h])$. Hence, another way to calculate $P(H(x) = h)$ is to use the below equation:

$$\begin{aligned} P(H(x) = h) &= P(D[h-1] < x \leq D[h]) \\ &= \int_{z_1=x}^{d^{max}[h]} \int_{z_2=d^{min}[h-1]}^x \mathfrak{d}^{h,h-1}(z_1, z_2) dz_2 dz_1. \end{aligned} \quad (6.42)$$

The discrete version of the above formula is presented below:

$$\begin{aligned} P(H_D[m] = h) &= P(D_D[h-1] < m \leq D_D[h]) \\ &= \sum_{n_1=m}^{d_D^{max}[h]} \sum_{n_2=d_D^{min}[h-1]}^{m-1} \mathbf{D}^{h,h-1}[n_1, n_2]. \end{aligned} \quad (6.43)$$

6.5 Numerical Results

In this section, we first describe the model we use for simulations, then we present the results.

6.5.1 Model Used for Simulations

Here we describe the two transmission range assignments we use for simulations, and we prove that these assignments are progressive. We used the discrete version of the formulas to compare analysis and simulation results. The two transmission range assignments we consider are identical transmission range assignment, which is used widely in the literature [1, 45, 48, 49], and the proposed transmission range assignment described in Chapter 5.

Theorem 14. *Both the identical transmission range assignment and the transmission range assignment with constant number of covered nodes are progressive.*

Proof. For the identical transmission range assignment, since $R(x) = r = cte, \forall x \in [0, L]$ it is obvious that it is a continuous function, and if $x < x'$ we have $R(x) + x = r + x < r + x' = R(x') + x'$, hence this assignment is progressive.

For the assignment defined by Equation (5.3), we note that since the integral of a continuous function is a continuous function, the transmission range assignment with constant number of covered nodes, which is the integral over some interval of $\lambda(x)$ is continuous. To prove the property $x' + R(x') > x + R(x)$ for $0 \leq x < x' \leq L$, we assume that there are no parts in the network with density function equal to zero. If some parts with $\lambda(x) = 0$ exist, there will be no nodes in them, and they do not have any effect on carrying the data towards the destination. Hence we can remove them at the first place, knowing that having them removed does not affect the network properties. Assuming $\lambda(x) \neq 0$, the density function is a continuous positive function, hence the integral of it is a strictly increasing function. Therefore the function $x + R(x)$ is strictly increasing, i.e., we have $x' + R(x') > x + R(x)$ for $0 \leq x < x' \leq L$.

□

6.5.2 Results

We verify our analysis by simulating 100,000 network realizations using the two transmission range assignments described in Section 6.5.1. To generate network realizations, we use the occupancy probabilities of the segments. In a network with M segments, we consider a node in segment m , according to a Bernoulli random variable with the probability of success equal to $P_{OC}[m]$ ³. Also, we add two nodes at the two ends of the network representing the source and the destination.

In our simulations, we select $\Delta x = 10$ cm, and $L = 700$ meters, which results in $M = 7,000$ segments. The network nodes are distributed based on the density function shown in Fig. 5.4, which is a real-world density function given in Example 4 of [1]. The probability of connectivity for the two schemes, when we have $K = 6$, is given in Fig. 6.2. The figure shows that our analysis for calculating the probability of connectivity using Equation (6.15) is confirmed by the simulations.

To verify the analysis of the hop-count random variable for any location of the network, $H_D[m]$, we present the derived pmf of this random variable for two segments 4,000 and 6,000 (which correspond to locations 400 and 600 meters from the source) along with the results obtained by simulations in Figs. 6.3 and 6.4, respectively. These results are for the networks with the proposed transmission range assignment. We see that the simulation results confirm the correctness of our analysis in both figures. Also, as we pick a location further away from the source, the number of hops needed for the data to reach it will be larger.

The probability for the locations of the network to receive the data by less than or equal to a certain number of hops are shown in Figs. 6.5 and 6.6 for the identical transmission range assignment and the proposed transmission range assignment, respectively. The different curves present different number of maximum allowable hops. This probability measure can be used to determine the probability that a specific location receives the data from the source within a limited time delay. Based on Equation (6.26), the envelope of the curves in each figure gives the probability of connectivity of the locations of the network to the source, for that assignment. As

³Network realizations can be obtained by the continuous concepts too. First, The number of the nodes in the network, N , is acquiring using Poisson distribution with mean $\bar{\lambda} = \int_{x=0}^L \lambda(x)dx$. Then, given the number of nodes, each node is independently distributed according to a known pdf $f(x) = \lambda(x)/\bar{\lambda}$. The resulting network has non-homogeneous Poisson distributed nodes [73].

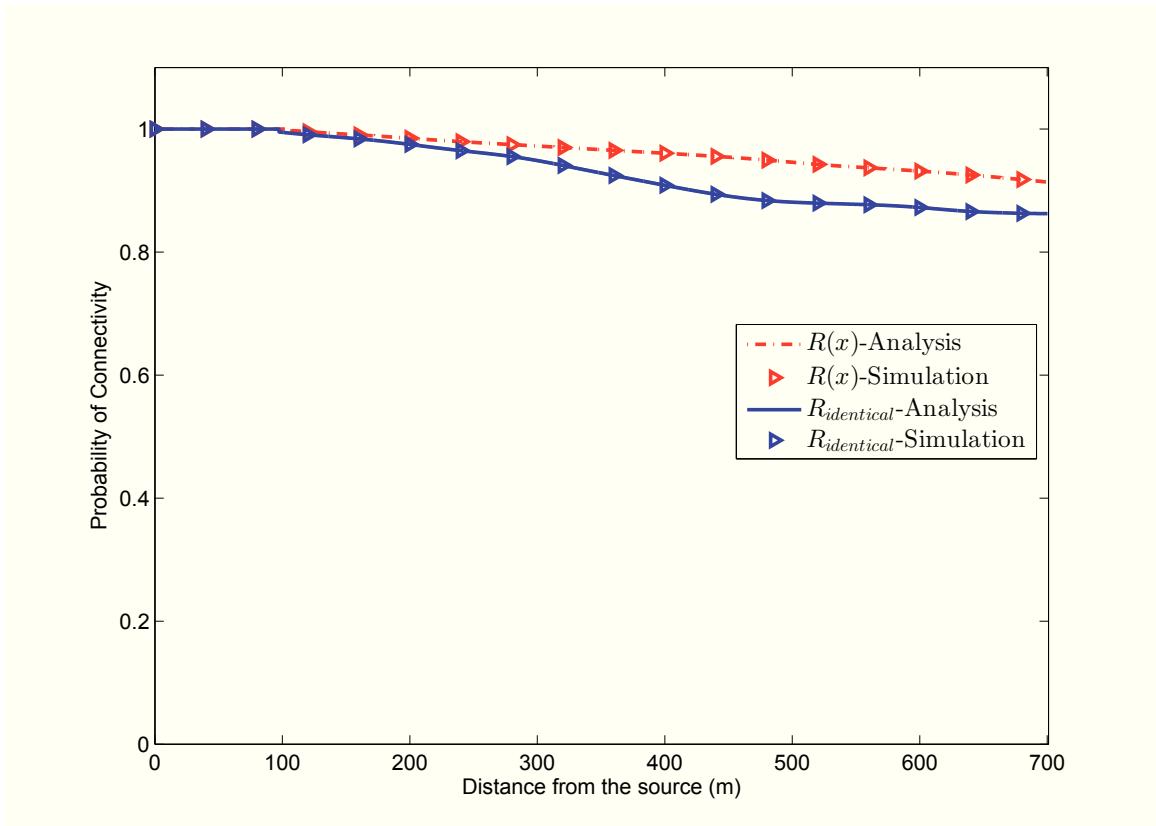


Figure 6.2: Probability of connectivity of the network with the given density function in Fig. 5.4 obtained by our analysis and simulation for the identical and the proposed transmission range assignments ($\mathcal{K} = 6$).

the figures show, the proposed transmission range assignment broadcasts with higher probability of connectivity while a lower number of hops is needed for the data to be delivered to different locations of the network.

Finally, verification of the data dissemination analysis, based on Equation (6.41), is done by simulation, and the results are given in Fig. 6.7. The curves in Fig. 6.7 show the probability of the last reachable location by at most a certain number of hops (indicated in the figure) be greater than or equal to the segment on the x -axis.

The analysis are confirmed by simulations. Furthermore, the curves show that the networks with our proposed transmission range assignment carry the data to a location further away from the source in a certain number of hops, compared to the

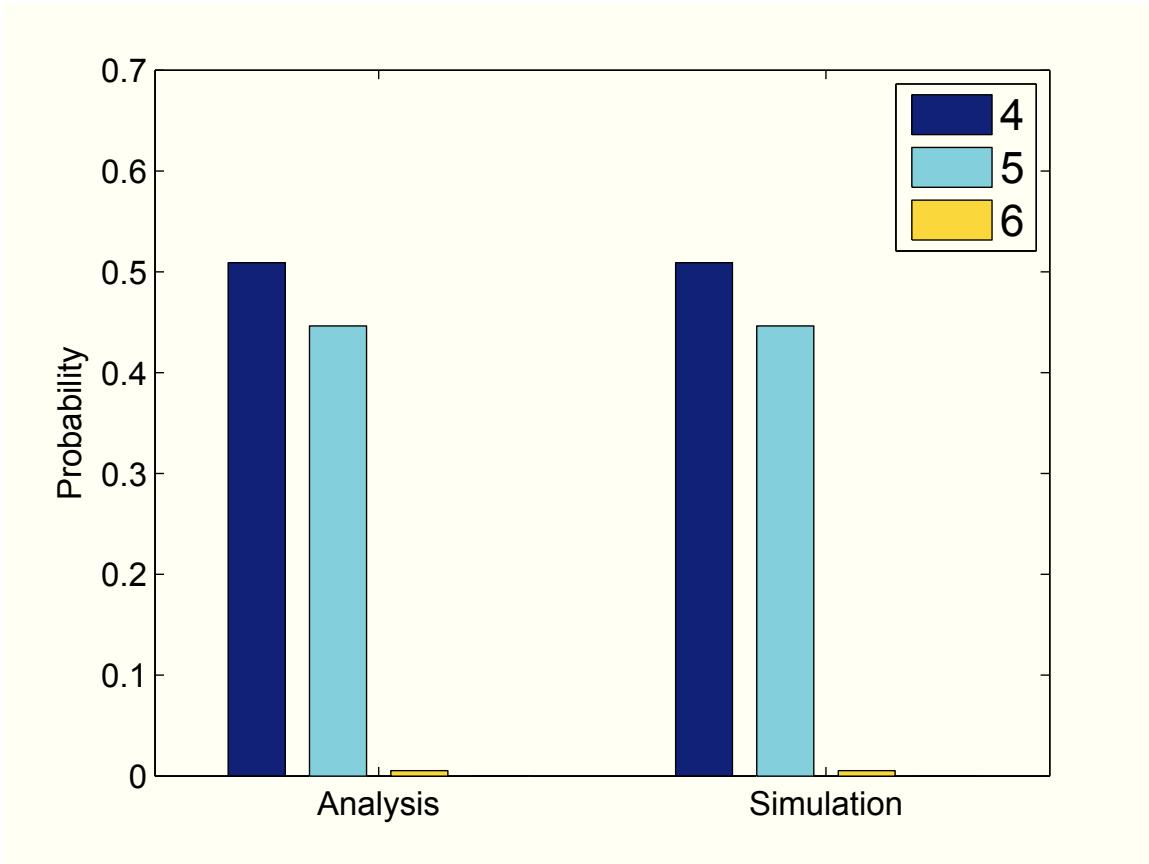


Figure 6.3: Hop count pmf of segment 4,000 (400 meters away from the source), obtained by analysis and simulation (for the proposed transmission range assignment).

identical transmission range assignment. Hence the proposed transmission range assignment is a better choice for the networks in which the data dissemination delay needs to be minimized, e.g., VANETs that want to broadcast safety messages as soon as possible.

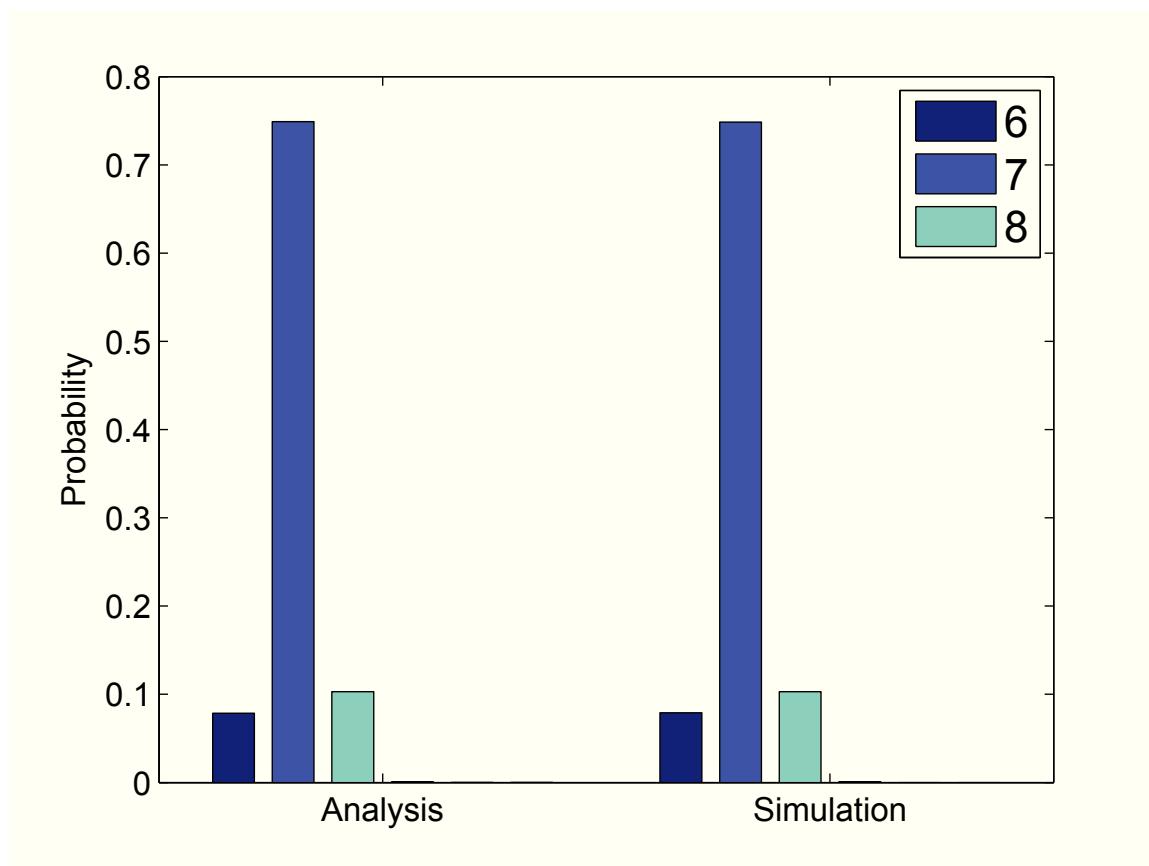


Figure 6.4: Hop count pmf of segment 6,000 (600 meters away from the source), obtained by analysis and simulation (for the proposed transmission range assignment).

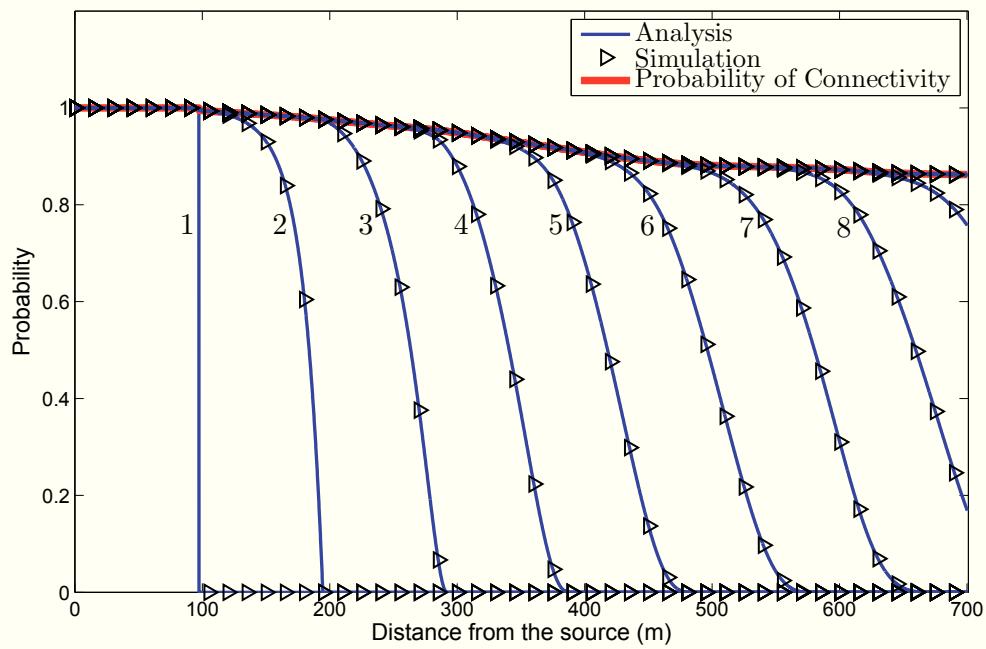


Figure 6.5: The data reception probability up to a certain number of hops for the identical transmission range assignment.

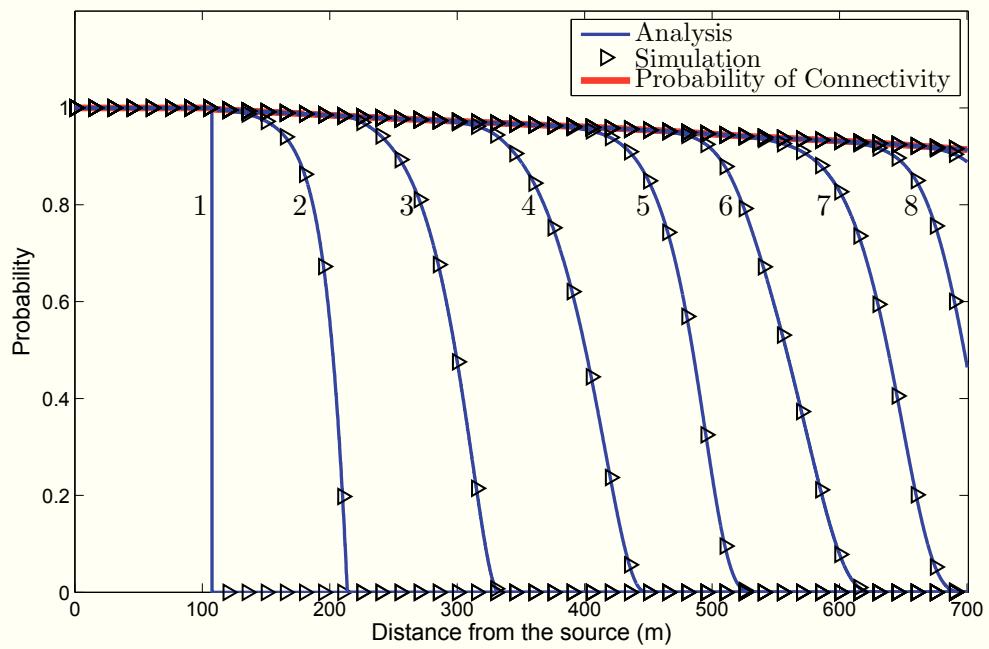


Figure 6.6: The data reception probability up to a certain number of hops for the proposed transmission range assignment.

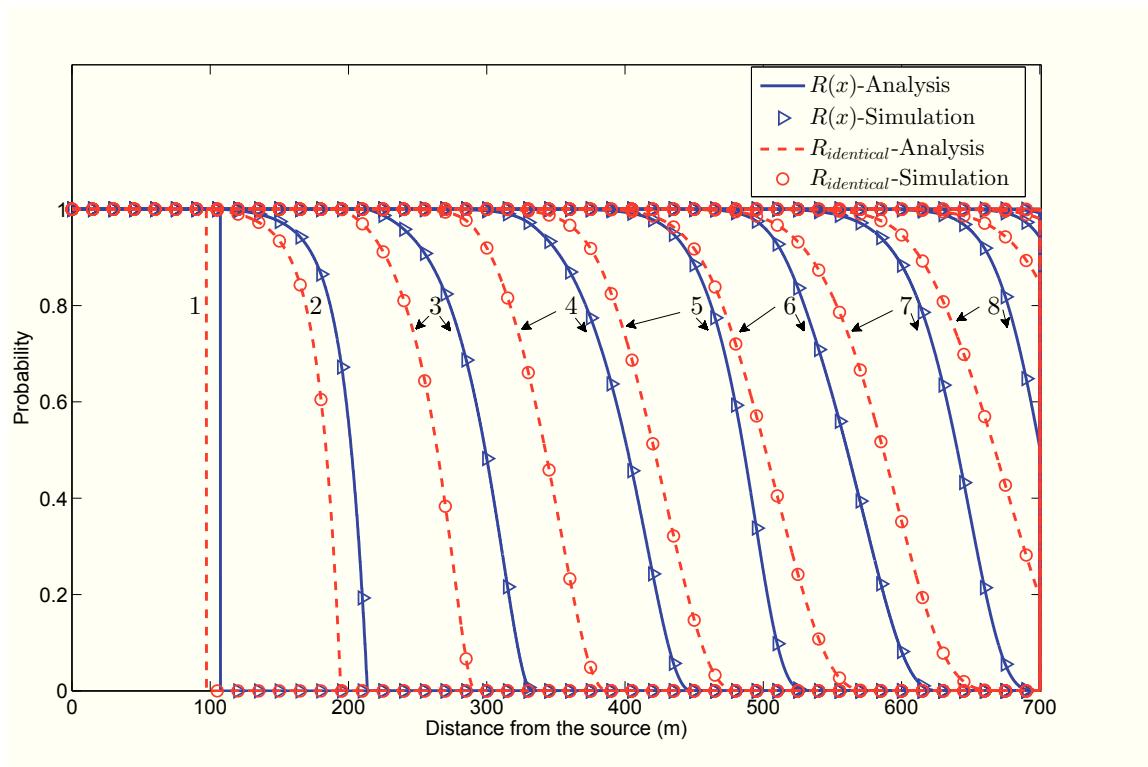


Figure 6.7: Probability for the last reachable location by at most h hops (h values are indicated in the figure) be greater than or equal to the segment on the x -axis (using Equation (6.41)).

Chapter 7

Localization and Location Verification

In this chapter, we use the derived formulas for the hop-count properties of one-dimensional wireless ad-hoc networks in Chapter 6, and we express the pmf of two hop-count random variables: the number of hops needed for a node located at an arbitrary location in the network to receive a message from a node located at one end of the linear network, and the number of hops needed for a node located at one end of the network to receive a message from a node at an arbitrary location. Based on the derived formulas, we then propose localization and location verification methods. Through simulations, we show that our proposed localization method not only has a competitive performance for a range-free method, but also outperforms range-based methods with a local distance measurement error of 10% or more. Furthermore, the proposed location verification protocol is shown to have better results compared to the existing verification systems that also use the hop-count information. An important feature of our methods is that they are applicable to arbitrary densities. This is unlike the existing methods that are limited only to the case of uniform node densities. Using simulations, we also evaluate the proposed schemes in the presence of Rician fading and show that their performance is rather robust with respect to the change in the fading parameter. Moreover, the hop-count equations derived in this work can be used in analyzing other aspects of broadcasting protocols such as quality of service and delay. After describing the system model and expressing the hop-count analysis for localization and location verification, we explain the use of our analysis through describing the proposed approach for localization and location verification. Finally, we compare our work with some of the existing methods for localization and location verification. The contents of this chapter have been published as [76, 77].

7.1 System Model

In this chapter again, the considered system is exactly the same as what we have in Chapter 5. For localization analysis, we assume that the anchors are fixed at locations L apart. For any location, only the information from the two nearest anchors (or just from one of them) is used. For location verification analysis, the verifiers are fixed at locations L apart, and any claimed location is verified by the two nearest verifiers (or just one of them). Hence, we just study the locations $0 \leq x \leq L$.

In both cases, we have two anchors/verifiers at two ends of the network. We later in the chapter show that the effect of considering the two anchors/verifiers for both the localization and location verification can be obtained by analyzing each anchor/verifier independently. Therefore, we study a network defined on $0 \leq x \leq L$ with one anchor/verifier at one end ($x = 0$). For localization, the flow of data is from the anchors towards the nodes, but in location verification, the flow is from the nodes towards the verifiers. Hence, we define two transmission range assignments $R_a(x)$ and $R_v(x)$ for any position $0 \leq x \leq L$, for localization and location verification purposes, respectively. Here again, we do not consider the effects of interference and collision caused by wireless communication among the nodes. It is thus assumed that the network is equipped with a proper multiple access control (MAC) scheduling to avoid interferences.

7.2 Hop-Count Analysis for Localization and Location Verification

In this section, we express two probabilities, based on the results obtained in Chapter 6. Here we just consider the continuous version of the derived formulas. First we express the probability of data reception in a certain number of hops for a specific location of the network from the anchor. Then we express the probability of data reception for the verifier from a specific location of the network in a certain number of hops. For each hop-count value, a group of nodes receives the data at the same time.

The number of hops required for a message sent by the anchor (the source node in the localization) to be received for the first time in location x is called the hop-count (from anchor) of location x and is denoted by $H_a(x)$. Using Equation (6.24), we have:

$$\mathbf{H}_a^{(h)}(x, y) = \int_{z=y_x}^y \lambda(z) \mathbf{H}_a^{(h-1)}(z, y_x) e^{-\int_{u=y_x}^z \lambda(u) du} dz,$$

for $y_x \leq y < x$. (7.1)

For any location $0 < x \leq L$, we have $P(H_a(x) = h) = \mathbf{H}_a^{(h)}(x, x)$, and thus

$$P(H_a(x) = h) = \int_{z=y_x}^x \lambda(z) \mathbf{H}_a^{(h-1)}(z, y_x) e^{-\int_{u=y_x}^z \lambda(u) du} dz,$$

for $0 < x \leq L$, (7.2)

where $\mathbf{H}_a^{(h)}(x, y)$ is derived recursively from Equation (7.1), with the initial condition $\mathbf{H}_a^{(0)}(0, 0) = 1$.

Now we are interested in finding the hop-count of the verifier located at location 0, when the source node is a node at location x that transmits a message towards the verifier. In this set-up, the nodes use the transmission range $R_v(\cdot)$. For any locations x and y , $x - R_v(x) \leq y < x \leq L$, we define $\mathbf{H}_v^{(h)}(x, y)$ to be the probability that the event “at least one node exists in the interval $[x - R_v(x), y]$ that sends the message to the verifier with hop-count $h - 1$, and no node that send the message with hop-count less than $h - 1$ to the verifier exists in that interval” happens. Using the progressive property of the transmission range $R_v(\cdot)$, and following the same steps as the ones in previous chapter, we derive

$$\mathbf{H}_v^{(h)}(x, y) = \int_{z=x-R_v(x)}^y \lambda(z) \mathbf{H}_v^{(h-1)}(z, x - R_v(x)) e^{-\int_{u=x-R_v(x)}^z \lambda(u) du} dz,$$

for $x - R_v(x) \leq y < x \leq L$. (7.3)

Also, for any location $0 < x \leq L$, we have $P(H_v(x) = h) = \mathbf{H}_v^{(h)}(x, x)$, and thus

$$\begin{aligned} \Pr(H_v(x) = h) &= \int_{z=x-R_v(x)}^x \lambda(z) \mathbf{H}_v^{(h-1)}(z, x - R_v(x)) e^{-\int_z^x \lambda(u) du} dz, \\ &\text{for } 0 < x \leq L, \end{aligned} \quad (7.4)$$

where $\mathbf{H}_v^{(h)}(x, y)$ is derived recursively using Equation (7.3) and the initial condition $\mathbf{H}_v^{(0)}(x, x) = 1$ for $0 < x \leq L$.

7.3 Localization

Using the obtained pmf's in Equation (7.2), we derive the pdf of the location of a node, given that it receives the message (sent by the anchor) with a certain number of hops. Based on this pdf, we design a protocol for the nodes to estimate their location. In our proposed protocol, the anchor broadcasts a message and the nodes observe their hop-count, and read their location from a table \mathfrak{T} based on their hop-count. The table \mathfrak{T} is calculated beforehand in a centralized fashion, and is broadcasted (e.g., as the message to be used for obtaining the hop-counts) to the nodes. This relieves the computational burden of the nodes, as there is one central unit that handles the calculations, and knows all the required information (i.e., the nodes density and the hop-count pmf's). As the nodes do not know yet their locations to use a variable transmission range assignment, we have to use an identical transmission range assignment for our localization protocol.

Define \mathbf{X} to be the random variable denoting the location of the node under consideration (from 0 to L). According to Equation (11.10) of [73], the distribution of the location of a node placed based on a non-homogeneous Poisson process with density $\lambda(x)$, assuming no additional knowledge, is:

$$f_{\mathbf{X}}(x) = \frac{\lambda(x)}{\int_0^L \lambda(x') dx'}, \quad \text{for } 0 \leq x \leq L. \quad (7.5)$$

For a node that receives the message by h hops, we are interested in finding the pdf $f_{\mathbf{X}|H_a(\mathbf{X})}(x|H_a(\mathbf{X}) = h)$, which gives the conditional pdf of the location of a node receiving the message (sent from the anchor) in h hops. Since the location of interest

is a random variable, we use $H_a(\mathbf{X})$. This conditional pdf can be rewritten as:

$$f_{\mathbf{X}|H_a(\mathbf{X})}(x|H_a(\mathbf{X}) = h) = \frac{f_{H_a(\mathbf{X})|\mathbf{X}}(h|\mathbf{X} = x)f_{\mathbf{X}}(x)}{f_{H_a(\mathbf{X})}(h)}. \quad (7.6)$$

In Equation (7.6), the term $f_{H_a(\mathbf{X})|\mathbf{X}}(h|\mathbf{X} = x)$ is equal to the known probability $P(H_a(x) = h)$, given by Equation (7.2), and $f_{\mathbf{X}}(x)$ is given by (7.5). The term $f_{H_a(\mathbf{X})}(h)$ in (7.6) is calculated as follows:

$$\begin{aligned} f_{H_a(\mathbf{X})}(h) &= \int_{x=0}^L f_{H_a(\mathbf{X})|\mathbf{X}}(h|\mathbf{X} = x)f_{\mathbf{X}}(x)dx \\ &= \int_{x=0}^L P(H_a(x) = h) \frac{\lambda(x)}{\int_0^L \lambda(x')dx'} dx. \end{aligned} \quad (7.7)$$

Finally, by using (7.7) in (7.6), we have:

$$f_{\mathbf{X}|H_a(\mathbf{X})}(x|H_a(\mathbf{X}) = h) = \frac{P(H_a(x) = h)\lambda(x)}{\int_{x'=0}^L P(H_a(x') = h)\lambda(x')dx'}. \quad (7.8)$$

Now we explain how to use the information of the other anchor (at the end of the network). The transmission range used for broadcasting from each anchor has to be identical throughout the network, but can be different from the transmission range of the other anchor. These transmission ranges are used for obtaining the formulas given in (7.2) and (7.4). Using the same approach as before, we have:

$$\begin{aligned} &f_{\mathbf{X}|H_a^{(1)}(\mathbf{X}), H_a^{(2)}(\mathbf{X})}(x|H_a^{(1)}(\mathbf{X}) = h_1, H_a^{(2)}(\mathbf{X}) = h_2) \\ &= \frac{f_{H_a^{(1)}(\mathbf{X}), H_a^{(2)}(\mathbf{X})|\mathbf{X}}(h_1, h_2|\mathbf{X} = x)f_{\mathbf{X}}(x)}{f_{H_a^{(1)}(\mathbf{X}), H_a^{(2)}(\mathbf{X})}(h_1, h_2)} \\ &= \frac{f_{H_a^{(1)}(\mathbf{X}), H_a^{(2)}(\mathbf{X})|\mathbf{X}}(h_1, h_2|\mathbf{X} = x)f_{\mathbf{X}}(x)}{\int_{x'=0}^L f_{H_a^{(1)}(\mathbf{X}), H_a^{(2)}(\mathbf{X})|\mathbf{X}}(h_1, h_2|\mathbf{X} = x')f_{\mathbf{X}}(x')dx'}, \end{aligned} \quad (7.9)$$

where random variables $H_a^{(1)}(\mathbf{X})$ and $H_a^{(2)}(\mathbf{X})$ represent the hop-count of the location random variable \mathbf{X} with respect to the first and the second anchor, respectively. Since the two anchors are located at the two ends of the network, the number of hops needed for the message sent from one anchor to reach a fixed location is independent of the number of hops needed for the message sent from the other anchor to reach that

location. Note that the two hop-counts are statistically dependent if no condition is imposed on the location in which the hop-counts are observed. For the scenario here, where the observation location is fixed, however, the two random variables are independent. For a sketch of the proof, assume that we cut the network into two different networks from the given fixed location. For each of the new networks, the number of hops needed for the data to traverse through the network, from the source (the anchor node at one end of the new network) to the other end of the new network (the given location in the original network) is independent from that of the other network. To see the independency more clear, using superscripts (1) and (2) for the data flow from the first and the second anchors, respectively, we have:

$$\begin{aligned}
& f_{H_a^{(1)}(\mathbf{X}), H_a^{(2)}(\mathbf{X})|\mathbf{X}}(h_1, h_2 | \mathbf{X} = x) = P(H_a^{(1)}(x) = h_1, H_a^{(2)}(x) = h_2) \\
&= \int \int_{z^{(1)}=y_x^{(1)}, z^{(2)}=y_x^{(2)}}^x \left(\lambda(z^{(1)}) \lambda(z^{(2)}) \mathbf{H}_a^{(1)(h_1-1)}(z^{(1)}, y_x^{(1)}) \mathbf{H}_a^{(2)(h_2-1)}(z^{(2)}, y_x^{(2)}) \right. \\
&\quad \left. e^{-\int_{u=y_x^{(1)}}^{z^{(1)}} \lambda(u) du} e^{-\int_{u=y_x^{(2)}}^{z^{(2)}} \lambda(u) du} dz^{(2)} dz^{(1)} \right) \\
&= \int_{z^{(1)}=y_x^{(1)}}^x \left(\lambda(z^{(1)}) \mathbf{H}_a^{(1)(h_1-1)}(z^{(1)}, y_x^{(1)}) e^{-\int_{u=y_x^{(1)}}^{z^{(1)}} \lambda(u) du} dz^{(1)} \right) \\
&\quad \times \int_{z^{(2)}=y_x^{(2)}}^x \left(\lambda(z^{(2)}) \mathbf{H}_a^{(2)(h_2-1)}(z^{(2)}, y_x^{(2)}) e^{-\int_{u=y_x^{(2)}}^{z^{(2)}} \lambda(u) du} dz^{(2)} \right) \\
&= P(H_a^{(1)}(x) = h_1) P(H_a^{(2)}(x) = h_2).
\end{aligned}$$

We thus have:

$$\begin{aligned}
& f_{\mathbf{X}|H_a^{(1)}(\mathbf{X}), H_a^{(2)}(\mathbf{X})}(x | H_a^{(1)}(\mathbf{X}) = h_1, H_a^{(2)}(\mathbf{X}) = h_2) \tag{7.10} \\
&= \frac{P(H_a^{(1)}(x) = h_1) P(H_a^{(2)}(x) = h_2) \lambda(x)}{\int_{x'=0}^L P(H_a^{(1)}(x') = h_1) P(H_a^{(2)}(x') = h_2) \lambda(x') dx'}.
\end{aligned}$$

Our calculations result in finding a pdf for the location of the node under consideration. Having the pdf, we can use the minimum mean-square error (MMSE) or the maximum a posterior (MAP) method to estimate the location of the node. Here, we use MMSE and find the expected value of the pdf of the location of the node that receives the message with h_1 and h_2 hops from the two anchors, respectively, and

store it in table $\mathfrak{T}(h_1, h_2)$. One of the anchors broadcasts this table to the nodes.

7.3.1 Enhancement

To further increase the estimation accuracy, we use the order of the nodes' locations from one end of the network to the other end. The closer a node to location 0, the lower order it has. First we discuss a general case in which we assume that a central unit knows the order of all the nodes, and it uses this information along with the hop-counts information of the nodes to estimate their location. This approach is computationally expensive. To reduce the complexity, we then propose a simplified version of this approach. In the simplified version, we assume that the nodes know their relative location among all the other nodes with the same hop-counts from the two anchors.

To find the relative location, a node can use directional antenna for signal reception. It then needs to count the number of nodes with the same hop-count on its left and right to find its relative order. The same relative information can be sent to the central unit for obtaining the total order of all the nodes using a procedure described in the following. The central unit arranges the nodes, starting from the nodes with hop-count one from the anchor on the left side of the network, and the maximum hop-count from the other anchor. Among these nodes, the node which has no nodes with the same hop-count on its left will be the first node. The rest of the nodes will be arranged in a similar manner. The central unit can thus order all the nodes based on the received information about the hop-count of the nodes and their relative location.

General Approach

Assume that there are N nodes distributed in the network based on pdf $f_{\mathbf{X}}(x)$ obtained in Equation (7.5). We denote the nodes by $n \in \{1, 2, \dots, N\}$, where n indicates the order of the nodes as well. So node 1 is the closest node to location 0, and node N is the closest node to location L . Node n receives the message sent from the first and the second anchor in $h_1^{(n)}$ and $h_2^{(n)}$ hops, respectively. Knowing the order of all the nodes, the best MMSE location estimate of the nodes is the expected value of the joint pdf of their locations. This joint pdf is defined in (7.11).

$$\begin{aligned} & f_{\mathbf{X}_1, \dots, \mathbf{X}_N}^*(x_1, \dots, x_N) \\ & \triangleq f(x_1, \dots, x_N | H_a^{(1)}(\mathbf{X}_1) = h_1^{(1)}, H_a^{(2)}(\mathbf{X}_1) = h_2^{(1)}, \dots, H_a^{(1)}(\mathbf{X}_N) = h_1^{(N)}, H_a^{(2)}(\mathbf{X}_N) = h_2^{(N)}). \end{aligned} \quad (7.11)$$

Thus the best estimate is:

$$\hat{\mathbf{x}} = (\hat{x}_1, \dots, \hat{x}_N) = \int_{x_1=0}^L \int_{x_2=x_1}^L \cdots \int_{x_N=x_{N-1}}^L (x_1, \dots, x_N) f_{\mathbf{X}_1, \dots, \mathbf{X}_N}^*(x_1, \dots, x_N) dx_N \cdots dx_1, \quad (7.12)$$

where $\hat{\mathbf{x}}$ denotes the vector of the estimated locations for the N nodes, and \hat{x}_n denotes the estimated location of node n .

This approach has a number of disadvantages. First of all it is centralized, and the central unit has to receive the ordering information from all the nodes and calculate the complex joint pdf of the locations of all the nodes. Furthermore, the N -fold integration is computationally complex. To overcome these problems, in the following, we introduce a simplified solution.

Simplified Approach

In this approach, every node is assumed to know its relative location among all the nodes with the same hop-counts. To simplify the calculations, in the rest of this section, we use the following notation:

$$f_{h_1, h_2}(x) \triangleq f_{\mathbf{X} | H_a^{(1)}(\mathbf{X}), H_a^{(2)}(\mathbf{X})}(x | H_a^{(1)}(\mathbf{X}) = h_1, H_a^{(2)}(\mathbf{X}) = h_2). \quad (7.13)$$

Here, similar to the general approach, the estimated locations are calculated at a central unit and are broadcasted to the nodes. We denote the number of nodes that receive the message in h_1 and h_2 hops from the two anchors by N_{h_1, h_2} . So for any set of values of h_1 , h_2 , and N_{h_1, h_2} , the estimated locations vector $\hat{\mathbf{x}}_{h_1, h_2} = (\hat{x}_1, \dots, \hat{x}_{N_{h_1, h_2}})$ is derived centrally and saved in the three dimensional table $\mathfrak{T}(h_1, h_2, N_{h_1, h_2})$. This table is then broadcasted to the nodes. Using its own values of h_1 and h_2 and N_{h_1, h_2} , and knowing its relative location within the set of N_{h_1, h_2} nodes with the same hop-counts, each node will then look up its location estimate from the table.

As all of the N_{h_1, h_2} nodes receive the message with the same hop-counts, all of them have the same location pdf. Since $x_1 < x_2 < \dots < x_{N_{h_1, h_2}}$, the joint pdf of the location of all of them is calculated as $\mathcal{C}_{N_{h_1, h_2}} \prod_{n=1}^{N_{h_1, h_2}} f_{h_1, h_2}(x_n)$, where $\mathcal{C}_{N_{h_1, h_2}}$ is the normalization factor and is calculated as follows:

$$\mathcal{C}_{N_{h_1, h_2}} \triangleq \int_{x_1=0}^L \int_{x_2=x_1}^L \cdots \int_{x_{N_{h_1, h_2}}=x_{N_{h_1, h_2}-1}}^L \prod_{n=1}^{N_{h_1, h_2}} f_{h_1, h_2}(x_n) dx_{N_{h_1, h_2}} \cdots dx_1. \quad (7.14)$$

Therefore we have:

$$\begin{aligned} \hat{\mathbf{x}}_{h_1, h_2} &= (\hat{x}_1, \dots, \hat{x}_{N_{h_1, h_2}}) \\ &= \mathcal{C}_{N_{h_1, h_2}} \int_{x_1=0}^L \int_{x_2=x_1}^L \cdots \int_{x_{N_{h_1, h_2}}=x_{N_{h_1, h_2}-1}}^L (x_1, \dots, x_{N_{h_1, h_2}}) \prod_{n=1}^{N_{h_1, h_2}} f_{h_1, h_2}(x_n) dx_{N_{h_1, h_2}} \cdots dx_1. \end{aligned} \quad (7.15)$$

Here, similar to the general approach, we have to perform a multiple fold integration. To simplify the calculations, we rewrite the above formula as in (7.16).

$$\begin{aligned} \hat{\mathbf{x}}_{h_1, h_2} &= (\hat{x}_1, \dots, \hat{x}_{N_{h_1, h_2}}) \\ &= \mathcal{C}_{N_{h_1, h_2}} \int_{x_1=0}^L f_{h_1, h_2}(x_1) \int_{x_2=x_1}^L f_{h_1, h_2}(x_2) \cdots \int_{x_{N_{h_1, h_2}}=x_{(N_{h_1, h_2}-1)}}^L f_{h_1, h_2}(x_{N_{h_1, h_2}})(x_1, \dots, x_{N_{h_1, h_2}}) dx_{N_{h_1, h_2}} \cdots dx_1. \end{aligned} \quad (7.16)$$

The formula in (7.16) can be expressed as a recursion. We first define the following set of recursive functions:

$$\begin{aligned} F_{h_1, h_2}^{(0)}(x) &\triangleq 1, & F_{h_1, h_2}^{(R)}(x) &\triangleq \int_{u=x}^L f_{h_1, h_2}(u) F_{h_1, h_2}^{(R-1)}(u) du \\ \text{for } R &= 1, \dots, N_{h_1, h_2}. \end{aligned} \quad (7.17)$$

Using these functions, we define the following set of recursive equations.

$$\begin{aligned}
E_{h_1, h_2}^{(1, R)}(x) &\triangleq \int_{u=x}^L u f_{h_1, h_2}(u) F_{h_1, h_2}^{(R-1)}(u) du, \\
E_{h_1, h_2}^{(L, R)}(x) &\triangleq \int_{u=x}^L f_{h_1, h_2}(u) E_{h_1, h_2}^{(L-1, R)}(u) du
\end{aligned} \tag{7.18}$$

for $R = 1, \dots, N_{h_1, h_2}$ and $L = N_{h_1, h_2} - R + 1$.

Finally, we have:

$$\begin{aligned}
\hat{\mathbf{x}}_{h_1, h_2} &= (\hat{x}_1, \dots, \hat{x}_{N_{h_1, h_2}}) \\
&= \mathcal{C}_{N_{h_1, h_2}}(E_{h_1, h_2}^{(1, N_{h_1, h_2})}(0), E_{h_1, h_2}^{(2, N_{h_1, h_2}-1)}(0), \dots, E_{h_1, h_2}^{(N_{h_1, h_2}, 1)}(0)).
\end{aligned} \tag{7.19}$$

7.4 Location Verification

In our proposed verification approach, a node sends its claimed location to the verifier. In every hop, the relay nodes increment the total hop-count of the message by one. The verifier gets the message along with the number of hops needed for the message to reach it. Using the hop-count pmf provided in Equation (7.4) the verifier calculates the plausibility that the claim is true. This is done by comparing the obtained plausibility with a set of classification thresholds to assess the trustworthiness of the claimed location. A single threshold level leads to a decision about acceptance or rejection of this location association.

For this approach to work, the transmission range assignment of the nodes has to be progressive, and it is not limited to just the identical transmission range assignment. Furthermore, more than one verifier can be employed to enhance the decision making process.

7.4.1 Plausibility Calculation

We use the same mapping from the probability values of receiving the message by the verifier in a certain number of hops to the plausibility of the claimed location as the one used in [13]. Assume that the message sent from a node is received by the verifier in h^* hops, and this node's claimed location is \check{x} . First, for taking into account the

relative position of the probability associated with h^* in the entire pmf of random variable $H_v(\check{x})$, we find the maximum probability value in the pmf.

$$P_{max}(\check{x}) \triangleq \max_{h \in \mathbb{N}} (\text{P}(H_v(\check{x}) = h)), \quad (7.20)$$

where \mathbb{N} denotes the set of natural numbers.

We consider the difference between $P_{max}(\check{x})$ and $\text{P}(H_v(\check{x}) = h^*)$ as a measure of plausibility of the claim. The larger this difference becomes, the less one should trust the claimed location. We define a *probability slack* function to show this difference, as shown below:

$$S(\check{x}, h^*) \triangleq P_{max}(\check{x}) - \text{P}(H_v(\check{x}) = h^*). \quad (7.21)$$

Finally, the *distrust* in the claimed location based on the observed number of hops is calculated by scaling the probability slack function, i.e.,

$$\frac{S(\check{x}, h^*)}{P_{max}(\check{x})}. \quad (7.22)$$

When we have just one verifier in the network, the final plausibility is considered as $1 - \frac{S(\check{x}, h^*)}{P_{max}(\check{x})}$, and is compared with a set of classification thresholds to assess the trustworthiness of the claimed location.

To get better results, we use two verifiers (one at location 0, and the other one at location L) for the nodes located in the interval $[0, L]$. To merge the distrust level of the two verifiers to calculate the final plausibility $\mathcal{P}(\check{x})$, we use the following formula:

$$\begin{aligned} & \mathcal{P}(\check{x}) \\ & \triangleq 1 - \frac{\frac{P_{max}^{(1)}(\check{x}) - \text{P}(H_v^{(1)}(\check{x}) = h^*)}{P_{max}^{(1)}(\check{x})} P_{max}^{(1)}(\check{x}) + \frac{P_{max}^{(2)}(\check{x}) - \text{P}(H_v^{(2)}(\check{x}) = h^*)}{P_{max}^{(2)}(\check{x})} P_{max}^{(2)}(\check{x})}{P_{max}^{(1)}(\check{x}) + P_{max}^{(2)}(\check{x})} \\ & = 1 - \frac{S^{(1)}(\check{x}, h^*) + S^{(2)}(\check{x}, h^*)}{P_{max}^{(1)}(\check{x}) + P_{max}^{(2)}(\check{x})}, \end{aligned} \quad (7.23)$$

where the superscripts denote the verifier numbers (1 for the verifier at location 0, and 2 for the other verifier).

In Equation (7.23), we used $P_{max}^{(\nu)}(\check{x})$ as a measure of the confidence in the opinion of verifier ν ($\nu \in \{1, 2\}$ here). Since the random variables $H_v^{(1)}(\check{x})$ and $H_v^{(2)}(\check{x})$ are

independent, their calculated plausibilities are independent as well. To consider all the observations of all the verifiers, the $P_{max}^{(\nu)}(\check{x})$ and $S^{(\nu)}(\check{x}, h^*)$ values of all verifiers ν are collected at a verification centre and a common plausibility $\mathcal{P}(\check{x})$ for the claimed location is calculated using Equation (7.23). This plausibility value is compared with a set of classification thresholds to assess the trustworthiness of the claimed location. In the simplest form, a single threshold level is used, and it leads to a decision about acceptance or rejection of this location association.

7.4.2 Performance Analysis

In contrast to most of the existing literature, we evaluate the performance analytically. The performance of a detection system is evaluated through its *probability of detection* and *probability of false alarm*. For a node located at (actual) location x^* , while its claimed location is \check{x} , we define two events: detection and false alarm. We assume that the claim is true if $|\check{x} - x^*| \leq d$, where d is a constant, and depends on the sensitivity of the system on the precision of nodes' locations. We consider a binary decision-making process where the plausibility is compared with a single threshold value, denoted by $0 \leq \tau \leq 1$, i.e., we accept a claim if the plausibility is above τ , and reject it, otherwise.

Definition 16 (Detection Event). *The incorrectness of a claim is detected, if $\mathcal{P}(\check{x}) < \tau$, while $|\check{x} - x^*| > d$.*

Definition 17 (False Alarm Event). *A correct claim is detected as incorrect (and the system produces a false alarm), if $\mathcal{P}(\check{x}) < \tau$, while $|x^* - \check{x}| \leq d$.*

For the example shown in Fig. 7.1 the system detects an incorrect claim for $\check{x} < x^* - d$ (interval Δ_L^D) and $\check{x} > x_\tau^R$ (interval Δ_R^D). The system generates a false alarm for $x^* - d < \check{x} < x_\tau^L$ (interval Δ^F).

Assuming a uniform distribution in the interval $[0, L]$ for the claimed location, the average probability of detection for a given location x^* is calculated as follows:

$$P_D(x^*) = \frac{1}{L} \int_{\check{x}=0}^L \Pi(\tau - \mathcal{P}(\check{x})) \Omega_d(\check{x} - x^*) d\check{x}, \quad (7.24)$$

where $\Pi(x)$ is the unit step function and $\Omega_d(x)$ is defined by

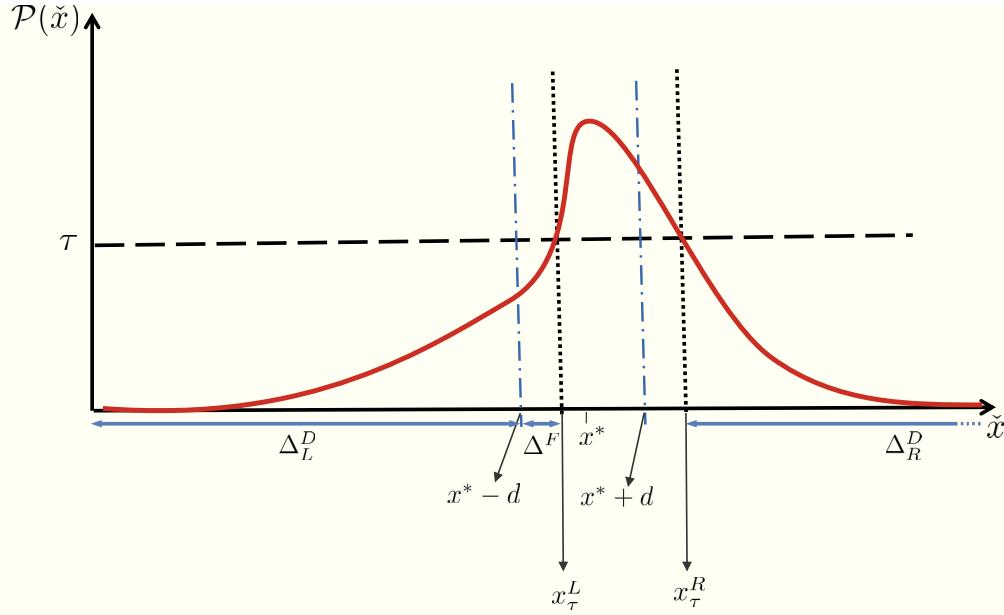


Figure 7.1: Detection happens for $\check{x} < x^* - d$ (interval Δ_L^D) and $\check{x} > x_\tau^R$ (interval Δ_R^D). False Alarm happens for $x^* - d < \check{x} < x_\tau^L$ (interval Δ^F).

$$\Omega_d(x) \triangleq \begin{cases} 1, & x < -d; \\ 0, & -d \leq x \leq d; \\ 1, & x > d. \end{cases} \quad (7.25)$$

The total probability of detection of the system is calculated using the pdf of x^* obtained in Equation (7.5), and is shown below:

$$\begin{aligned} \mathbb{P}_D &= \int_{x^*=0}^L f_{\mathbf{X}}(x^*) P_D(x^*) \\ &= \int_{x^*=0}^L \left(f_{\mathbf{X}}(x^*) \frac{1}{L} \int_{\check{x}=0}^L \Pi(\tau - \mathcal{P}(\check{x})) \Omega_d(\check{x} - x^*) d\check{x} \right) dx^*. \end{aligned} \quad (7.26)$$

Similarly, the probability of false alarm at a given x^* averaged over \check{x} is given by

$$P_F(x^*) = \frac{1}{L} \int_{\check{x}=0}^L \Pi(\tau - \mathcal{P}(\check{x}))(1 - \Omega_d(\check{x} - x^*)) d\check{x}, \quad (7.27)$$

and the total probability of false alarm by

$$\begin{aligned} \mathbb{P}_F &= \int_{x^*=0}^L f_{\mathbf{X}}(x^*) P_F(x^*) \\ &= \int_{x^*=0}^L \left(f_{\mathbf{X}}(x^*) \frac{1}{L} \int_{\check{x}=0}^L \Pi(\tau - \mathcal{P}(\check{x}))(1 - \Omega_d(\check{x} - x^*)) d\check{x} \right) dx^*. \end{aligned} \quad (7.28)$$

7.5 Numerical Results

Here, we first demonstrate that our analytical results match the Monte Carlo simulations. Then we compare our localization approach with some existing range-free and range-based approaches. Furthermore, we present some cases of the calculated location pdf's where the number of hops is given (derived in Equations (7.8) and (7.10)). Finally, we present simulation results related to location verification and compare our location verification approach with the work presented in [13].

In our simulations, we also consider the effects of fading. We use Rician fading model [78], which is appropriate for a linear network where a line-of-sight component is expected to exist in the received signal. Based on the results of [78], the parameter K of Rician fading for urban and semi-rural environments is between 1.8 to 6.6, with larger values of K corresponding to rural cases. We consider two cases of $K = 3$ and $K = 5$ for urban and semi-rural environments, respectively. To obtain our results, we simulated 100,000 network realizations¹ based on the density function shown in Fig. 5.4, which is a real-world density function given in Example 4 of [1]. Throughout the simulations, the transmission range used for localization is considered to be 70 meters for both anchors (i.e., $\forall x : R_a^{(1)}(x) = R_a^{(2)}(x) = 70$ meters). Also, the same range is used for $R_v^{(1)}(x)$ (to transmit the message towards verifier 1 at location 0) and $R_v^{(2)}(x)$ (to transmit the message towards verifier 2 at location $L = 700$ meters).²

¹In our simulations, we just considered connected networks, i.e., the cases where all the nodes are within the transmission range of another one.

²We fix the transmission power of the nodes to have a fixed reception power level (p_r) at the distance of 70 meters from the node in the no-fading environment. To consider the effects of fading,

Table 7.1: Hop-count pmf of the location $x = 450$ meters, receiving the message from the anchor at $x = 0$.

	$h = 6$	$h = 7$	$h = 8$	$h = 9$
Analysis	0	0.2724	0.6924	0.0350
No Fading	0	0.2717	0.6933	0.0349
Rician Fading (K=5)	0	0.3181	0.6514	0.0294
Rician Fading (K=3)	0.0114	0.3376	0.6267	0.0237

Table 7.2: Hop-count pmf of the verifier at $x = 0$, receiving the message from a node at location $x = 500$ meters.

	$h = 7$	$h = 8$	$h = 9$	$h = 10$
Analysis	0	0.4895	0.4929	0.0176
No Fading	0	0.4891	0.4933	0.0176
Rician Fading (K=5)	0.0132	0.5382	0.4335	0.0149
Rician Fading (K=3)	0.0418	0.5752	0.3705	0.0121

In the first row of Tables 7.1 and 7.2, we have presented the pmf's of $H_a(x)$ and $H_v(x)$, respectively. These pmf's are derived based on Equations (7.2) and (7.4), for $x = 450$ and $x = 500$ meters, respectively. In the same tables, we have also included the results of Monte Carlo simulations in the absence and presence of fading. As can be seen in both tables, the simulation results with no fading match well with our analysis. The tables also show that the introduction of fading changes the distributions, with larger deviation from the analysis for the smaller value of K ($K = 3$). This is expected, as the smaller the K value, the weaker the line-of-sight component of the received signal would be. A careful comparison of fading results with no fading ones, however, show a rather small difference in the pmf, even for the case of $K = 3$.

the nodes transmit with the same fixed power through a Rician channel, and we measure the distance from them that the signal is received with power p_r . The measured distance is the effective transmission range of the nodes.

7.5.1 Localization

In Fig. 7.2, we have shown the pdf of the location of a node that receives the message in $h = 5$ hops from an anchor at $x = 0$. The pdf is derived using Equation (7.8). The simulated pdf in the absence and presence of fading is also shown in Fig. 7.2. Again, while the simulated pdf in the absence of fading matches the analytical result, the presence of fading changes the pdf, with a larger change for the smaller value of K . Note that even for the case of $K = 3$, the MMSE location estimate (i.e., 300.4 meters) changes by only about 5% compared to the case with no fading (i.e., 285.1 meters).

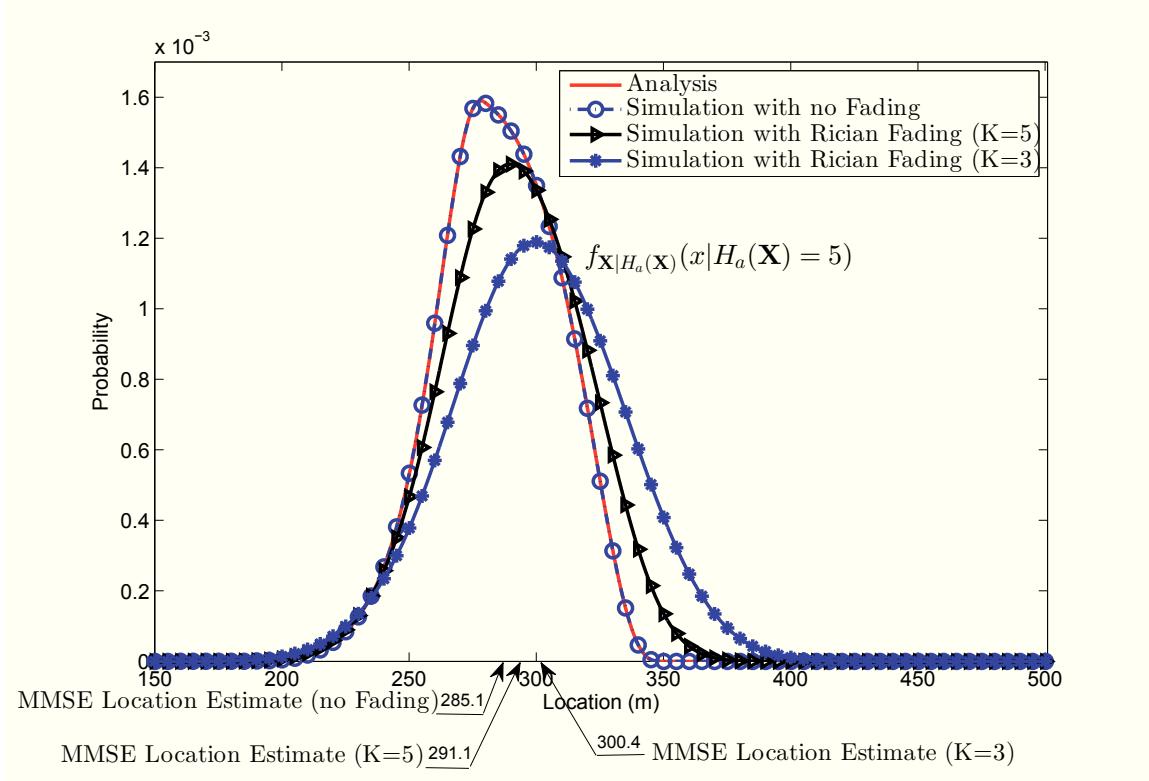


Figure 7.2: The pdf of the location of a node that receives the message by $h = 5$ hops from the anchor at $x = 0$.

Fig. 7.3 shows the location pdf for the two cases where there is one anchor at one end of the network, and where there are two anchors at two ends of the network (at locations 0 and 700 meters). The latter pdf is obtained using Equation (7.10). The node under consideration receives the message in $h_1 = 5$ hops from the first anchor

(at 0), and in $h_2 = 6$ hops from the other anchor. The comparison of the analysis with simulations in the absence of fading reveals a perfect match. As can be seen from Fig. 7.3, using two anchors results in a narrower pdf and consequently a better localization. Furthermore, using the joint pdf results in some cancellation of the shift in the location estimates, caused by fading, and thus having closer estimates to the case of no-fading compared with the situation where marginal pdf's are used. In fact, with two anchors, the MMSE location estimate in the presence of fading is only less than 1% different from that of the analysis.

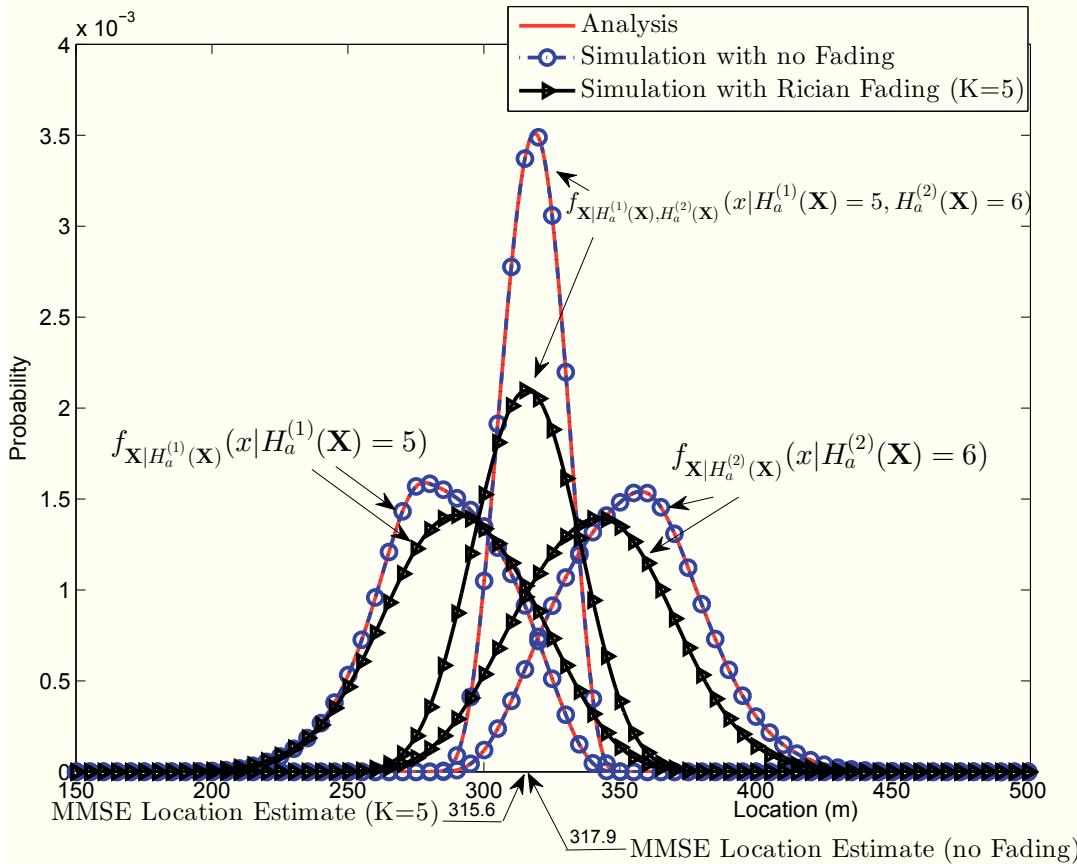


Figure 7.3: The pdf of the location of a node that receives the message by $h_1 = 5$ and $h_2 = 6$ hops from two anchors at $x = 0$ and $x = L = 700$ meters, respectively.

The comparison between our proposed methods and DV-Hop, DV-RND, and some hypothetical range-based approaches is shown in Fig. 7.4 in terms of the localization error normalized with respect to the transmission range. No fading is considered in these simulations. For the range-based approaches, we assume that a node estimates

its distance to a node that transmits to it, but with some measurement error. For simulations, the measurement errors are drawn from Gaussian distributions with standard deviation equal to $\epsilon\%$ of the actual distance. Here, we simulated the range-based approaches for $\epsilon \in \{5\%, 10\%, 15\%, 20\%, 25\%\}$. To see the effect of node density on the results, we consider 25 different density functions, each having the same general shape as that of Fig. 5.4, but multiplied by a different γ , where $\gamma \in \{0.2, 0.4, \dots, 5\}$. For each (connected) network realization, the difference between the actual location of a node and its location estimate is considered as the localization error of that node. The average error for all the nodes of a network is then considered as the localization error of that network. The overall average of the localization error for each γ value is then calculated as the average of the localization errors of 100,000 network realizations with that density.

As can be seen from the figure, the performance of all approaches improves with the increase in node density. Our proposed localization approach with two anchors outperforms the DV-Hop approach for a wide range of sufficiently large γ values. Fig. 7.4 also demonstrates that the performance of our enhanced approach improves with γ at a rate considerably faster than that of the other approaches. With high density values, the probability of having more nodes with the same pdf grows. The more nodes have the same location pdf (for certain hop-counts from two anchors), the more expected values will be obtained from the enhanced approach's formulas. This results in having less distance between these expected values, which, on average, results in less localization error.

Due to the above reasoning, for very large values of γ , the localization error of the enhanced approach tends to zero. This is while the performance of other approaches (except for DV-RND) saturates with increasing γ , as shown in Fig. 7.4. The original proposed approach and the DV-Hop approach use one (expected) value for each pdf (for any pair of hop-counts) as the location estimate of the nodes with the same hop-counts. Hence, the location error converges to the standard deviation of the location pdf, and thus to a constant value for both the original proposed approach and the DV-Hop approach. For the DV-RND approach, the larger the number of nodes that provide RND (as a result of larger density), the better the location estimate for all the nodes. This comes with the cost of more RND exchange, hence more overhead. For the considered range-based approaches, with the increase of γ , the localization error converges to the standard deviation of the sum of some Gaussian random variables

ranging from (near) 0 to R_a .

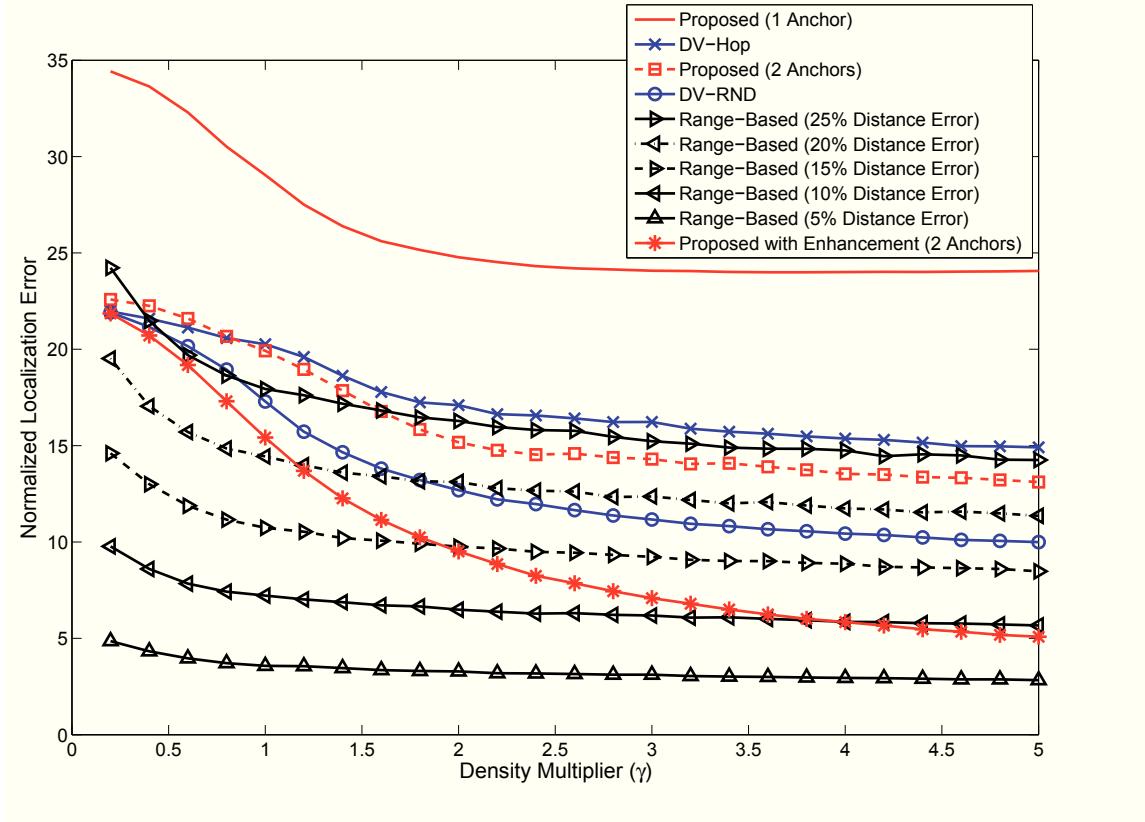


Figure 7.4: Localization error of different approaches for different densities, while $\forall x : R_a^{(1)}(x) = R_a^{(2)}(x) = 70$ meters.

The enhanced proposed approach outperforms DV-RND for the whole range of γ values, and even the range-based approaches with local distance measurement error as low as 10% of the actual distance for values of γ larger than 4. As we pointed out earlier, it is shown that the distance measurement error often exceeds 10% [59, 60], sometimes quite significantly so.

Some elements of the localization table \mathfrak{T} for $N_{h_1, h_2} = 3$, used for the networks with $\gamma = 1$, are given in Table 7.3. These values are used as the location estimates for nodes within the set of 3 nodes with the same hop-counts. The larger the value of h_1 , the farther away the location estimate from the location 0. This is while for larger values of h_2 , the location estimate gets closer to location 0.

To show an example of the values of the localization table \mathfrak{T} for different number

Table 7.3: Some elements of the localization table \mathfrak{T} with $N_{h_1, h_2} = 3$, for $\gamma = 1$.

\dots	$h_2 = 6$	$h_2 = 7$	$h_2 = 8$	$h_2 = 9$	\dots
$h_1 = 6$	$\begin{pmatrix} 334.5 \\ 350.1 \\ 365.7 \end{pmatrix}$	$\begin{pmatrix} 305.7 \\ 321.7 \\ 337.8 \end{pmatrix}$	$\begin{pmatrix} 280.9 \\ 298.1 \\ 315.2 \end{pmatrix}$	$\begin{pmatrix} 260.0 \\ 277.2 \\ 294.2 \end{pmatrix}$	
$h_1 = 7$	$\begin{pmatrix} 362.5 \\ 378.7 \\ 394.9 \end{pmatrix}$	$\begin{pmatrix} 332.6 \\ 350.1 \\ 367.6 \end{pmatrix}$	$\begin{pmatrix} 308.5 \\ 325.6 \\ 342.7 \end{pmatrix}$	$\begin{pmatrix} 289.1 \\ 304.3 \\ 319.9 \end{pmatrix}$	
$h_1 = 8$	$\begin{pmatrix} 385.6 \\ 403.0 \\ 420.4 \end{pmatrix}$	$\begin{pmatrix} 357.6 \\ 374.8 \\ 392.0 \end{pmatrix}$	$\begin{pmatrix} 334.3 \\ 350.1 \\ 366.0 \end{pmatrix}$	$\begin{pmatrix} 315.1 \\ 328.0 \\ 341.8 \end{pmatrix}$	
$h_1 = 9$	$\begin{pmatrix} 407.2 \\ 424.3 \\ 441.4 \end{pmatrix}$	$\begin{pmatrix} 380.7 \\ 396.4 \\ 411.9 \end{pmatrix}$	$\begin{pmatrix} 358.5 \\ 372.2 \\ 385.2 \end{pmatrix}$	$\begin{pmatrix} 337.7 \\ 350.1 \\ 362.5 \end{pmatrix}$	
\dots					

of nodes with the same hop-counts, we present the values of this table for different values of N_{h_1, h_2} , when $h_1 = 5$ and $h_2 = 6$, in Table 7.4. These values are also for networks with $\gamma = 1$.

7.5.2 Location Verification

The obtained plausibility $\mathcal{P}(\tilde{x})$ for one verifier and two verifiers (from Equation (7.23)) are shown in Fig. 7.5. In the same figure, we show the plausibility of the PLV approach [13], which we call the original PLV approach here. The original PLV approach assumes that the nodes are distributed uniformly in the network. For this approach to work, we took the average of the density function (i.e., $\bar{\lambda} \triangleq \frac{1}{L} \int_{x=0}^L \lambda(x)$), and fed it to the original PLV as the (uniform) density of the network. The illustrated plausibilities are shown for locations in the interval [300, 450] meters, and are for a node located at (actual) location $x = 375$ meters. Generally, a higher number of verifiers can detect a false location claim with a higher probability. The plausibility of our approach when we have two verifiers in the network is narrower than the same

Table 7.4: Some elements of the localization table \mathfrak{T} for different values of N_{h_1,h_2} , (for $h_1 = 5$, $h_2 = 6$ and $\gamma = 1$).

$N_{h_1,h_2} = 1$	$N_{h_1,h_2} = 2$	$N_{h_1,h_2} = 3$	$N_{h_1,h_2} = 4$	$N_{h_1,h_2} = 5$	\dots
$\begin{pmatrix} 317.9 \end{pmatrix}$	$\begin{pmatrix} 311.9 \\ 323.9 \end{pmatrix}$	$\begin{pmatrix} 308.9 \\ 318.0 \\ 326.8 \end{pmatrix}$	$\begin{pmatrix} 306.9 \\ 314.7 \\ 321.2 \\ 328.7 \end{pmatrix}$	$\begin{pmatrix} 305.5 \\ 312.6 \\ 318.0 \\ 323.4 \\ 330.0 \end{pmatrix}$	

system using the original PLV approach. This results in more accurate detections. (For the case of one verifier, our approach and PLV perform closely.)

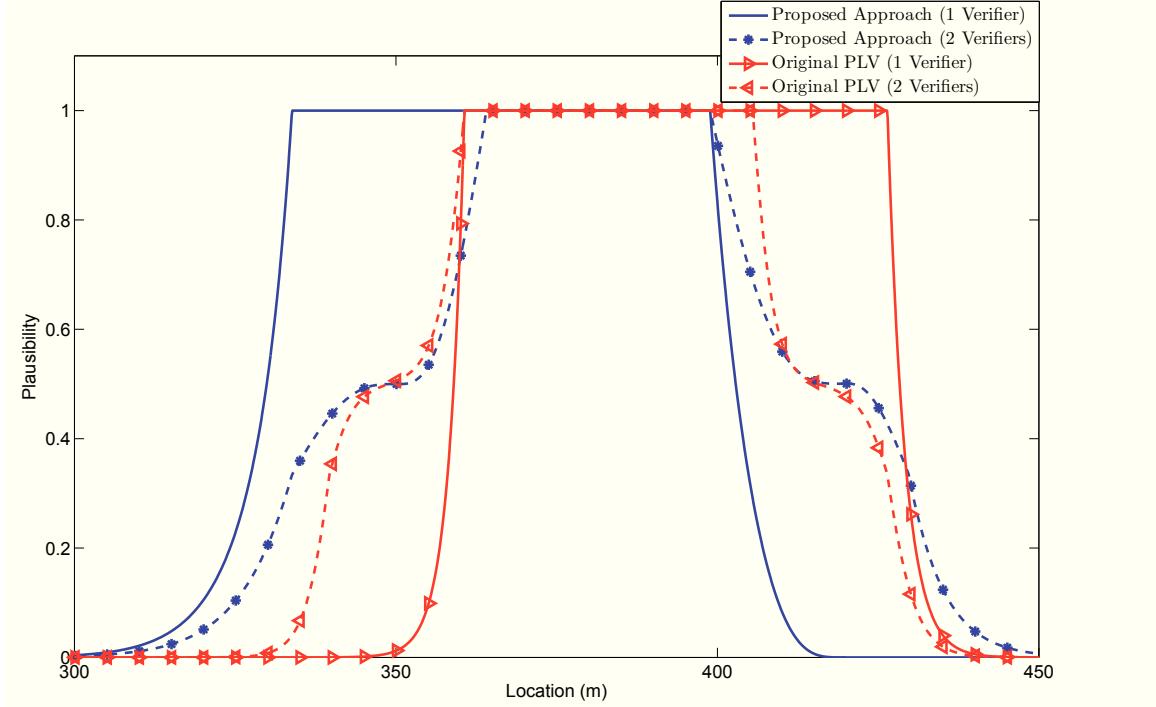


Figure 7.5: Plausibility of all the locations in the interval [300, 450] meters for different approaches, where the claim is from a node located at $x = 375$ meters.

An indication of the success of a classification method is the receiver operating characteristic (ROC) curve. This curve is the plot of the probability of detection as a function of the probability of false alarm of the system [79]. We generated ROCs

by sweeping the range of the classification threshold, τ , from 0 to 1 in increments of 0.01. A good classifier provides high detection probabilities for very small values of false alarm probability, i.e., the closer the ROC is to the point (0, 1), the better the classification performs. As expected, Fig. 7.6 shows that the system with two verifiers outperforms the system with one verifier. Also, it can be seen that our verification system has better ROC curves compared with the original PLV approach, when the two systems use the same number of verifiers. To better demonstrate the advantage of our scheme, we have zoomed into the important part of the ROC curves in the vicinity of the point (0, 1). As can be seen, for example, at $P_D = 0.92$, our scheme has $P_F = 0.005$ vs. 0.027 for the original PLV for the two verifier case.

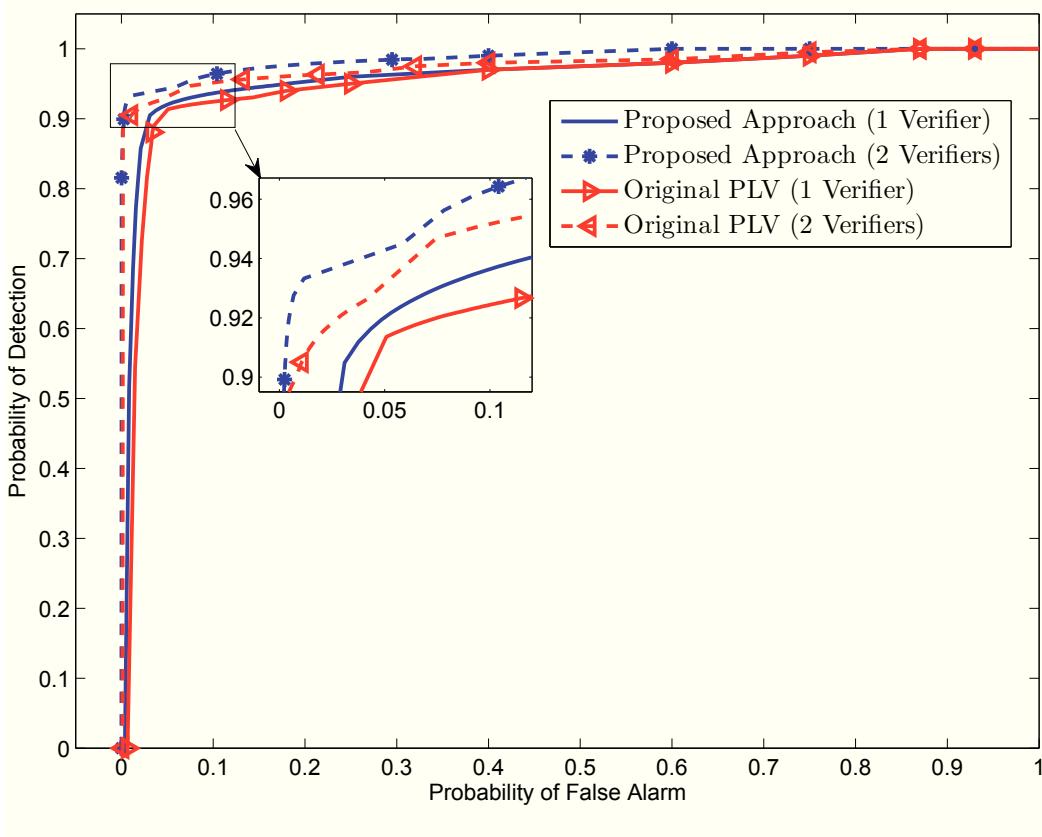


Figure 7.6: Receiver operating characteristic (ROC) curve of different approaches.

Chapter 8

Conclusion and Future Work

8.1 Conclusion

In this thesis, we considered wireless networks with known and unknown node locations. For the networks where the location of the nodes is known, we solved the Minimum-Energy Broadcasting problem for the linear (one-dimensional) cases (Chapter 3), as well as for cross networks (Chapter 4). Our solutions consist of optimal and some near-optimal algorithms. We also presented distributed algorithms which have lower time complexity compared with the centralized algorithms, while they perform close to the centralized ones. We demonstrated that the distributed algorithm performs well even for grid networks with perpendicular segments.

In Chapters 5, 6 and 7, we studied linear networks where the knowledge of the exact location of the nodes is not available, and the locations are known probabilistically. We proposed a transmission range assignment to make the network connected with high probability, while the average total consumed power is limited in Chapter 5. The proposed assignment is obtained by maintaining the same probability of having at least one node in the transmission range of the nodes located anywhere in the network. The proposed assignment achieves a higher probability of connectivity compared with an identical transmission range assignment, which is commonly used in the literature. Furthermore, it is shown that our approach, on average, has a bounded contention level, while this quantity can be very large for the identical assignment.

For the network model used in Chapter 5, we studied the probability of connectivity of any location of the network to the source, the hop-count probability mass function (pmf) of an arbitrary location of the network, and the pdf of the maximum coverage (last reachable distance from the source) for a given number of hops in Chapter 6.

The proposed analysis is applicable to progressive transmission range assignments, which include identical transmission range assignments. Analyzing these features helps to perform localization analysis, as well as location verification. Also, they can be helpful in energy consumption analysis.

In Chapter 7, based on the analysis derived in Chapter 6, we expressed exact formulas for the hop-count of any location of a one-dimensional wireless ad-hoc network when the sender is at a certain location, as well as the hop-count of a receiver at a certain location where the sender's location can be anywhere in the network. We proposed a localization and a location verification method based on the derived formulas. We validated our results via simulation, and compared them with some known approaches, with and without fading. The proposed localization method outperforms range-based methods with local distance measurement error of 10% and more, and state-of-the-art range-free localization protocols. The proposed location verification protocol was compared with the probabilistic location verification (PLV) method and was shown to perform better, e.g., by having a better receiver operating characteristic (ROC) curve.

8.2 Future Work

Here are some of the topics that can be studied further:

- Finding a transmission range assignment that maximizes the *average* probability of connectivity of a linear network with probabilistic node locations. We will present some preliminary results of our study on this topic in Subsection 8.2.1.
- Finding the transmission range assignment which consumes the lowest possible energy, while guaranteeing a certain probability of connectivity to the source along the network, for linear networks with probabilistic node locations.
- Extending the performed study to find optimal and near-optimal solutions for minimum-energy broadcasting on cross networks with non-perpendicular lines.
- Enhancing the accuracy of the derived formulas for localization in Chapter 7 in fading environments.
- Considering non-stationary node density functions, and cases in which the density is estimated with some error would be of practical interest.

- Applying the hop-count analysis results to various other problems, such as delay analysis (e.g., for broadcasting safety messages in VANETs), energy consumption analysis, and studying the quality of service in broadcasting protocols. One example would be to design a transmission range assignment to maximizing delay for a given amount of energy while achieving a minimum level of probability of connectivity. This range assignment can be used for applications such as: disseminating safety messages (to have minimum dissemination time), real-time applications (multimedia networks, etc.), power grids (to minimize the delay for the center to be notified of a malfunction in a power grid).

8.2.1 Average Probability of Connectivity

In Chapter 5, we proposed a heuristic transmission range assignment to maximize the probability of connectivity of the destination of a linear network to the source. The obtained formulas in that chapter and Chapter 6 enable us to study a new problem. Assume that there is no specific destination node to deliver the data to, and we are interested in delivering the data to as many nodes as possible, on average (as the nodes are distributed randomly on the line, and the exact location and the total number of the nodes are unknown) and with high probability. This type of transmission range assignments, to the best of our knowledge, has not been studied. One important application of these assignments is for safety message propagation. Consider a case where a car has an accident and wants to transmit this information to the other neighboring cars. It is obvious that the closer the neighboring cars, the more essential this information is for them. Here, there is no specific destination, and a good protocol maximizes the average probability of connectivity of the nodes to the source. It may be better to have a weighted average, i.e., the closer a node is to the source, the more weight it is given, but as a first step, one can just consider the average probability of connectivity (with identical weighing).

We define the goal of the desired transmission range assignment as maximization of the average probability of connectivity of the nodes to the source, and not just the probability of connectivity of the last destination. The average probability of connectivity up to location y is defined as following:

$$\overline{P^C}(y) = \int_{x=0}^y \lambda(x) P^C(x) dx. \quad (8.1)$$

The discrete version of the above formula is given below (the average probability of connectivity up to segment n):

$$\overline{P^C}[n] = \sum_{m=0}^n P_{OC}[m] P_D^C[m]. \quad (8.2)$$

To justify our proposed approach, we considered a linear network with uniformly distributed nodes on a line with total length $L = 1000$ meters. In our simulations, we discretize the network by selecting $\Delta x = 10$ cm, and we thus have $M = 10,000$ segments. Intuitively, the identical transmission range assignment is the best assignment employed for this kind of network. We compared the identical transmission range assignment with the three heuristic assignments shown in Fig. 8.1. All four assignments have the same average total consumed power (defined by Equation (5.7) for the discrete network model). Note that transmission range assignments $R_{\text{identical}}$ (the identical transmission range assignment), R_2 , and R_3 decay towards the end of the network to limit the range so that no power is wasted by communicating beyond the destination (R_1 does not cover beyond the destination, hence there is no need for decaying it).

As it can be seen in Fig. 8.2, the identical transmission range assignment has the maximum probability of connectivity of the destination to the source (using either Equations (5.8) or (6.15)).

For the average probability of connectivity, using Equation (8.2), as we see in Fig. 8.3, the identical assignment is not the best, and other assignments outperform it. For example assignment R_1 has the best average probability of connectivity for the whole network, as well as for any location of the network. For the safety message dissemination applications, R_1 is the best among the studied assignments, as it has high average probabilities of connectivity near the source.

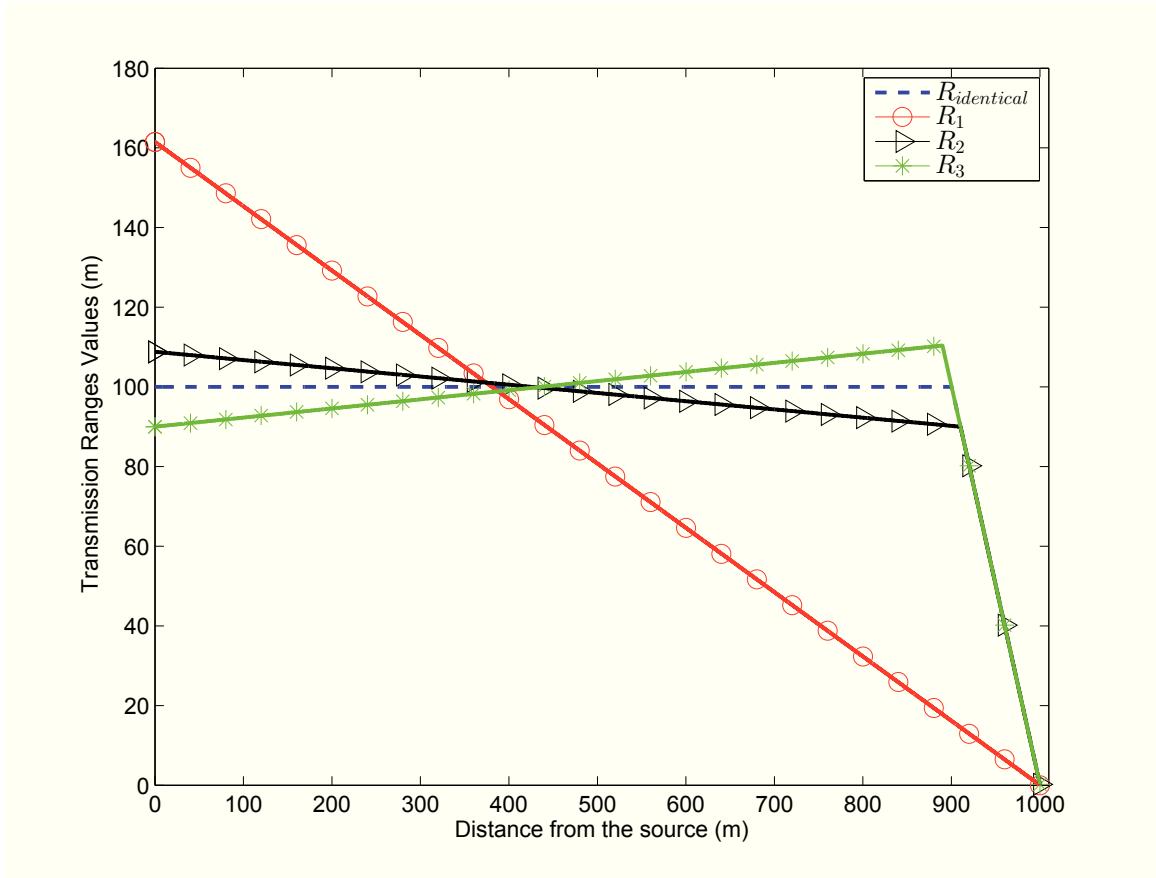


Figure 8.1: The identical transmission range assignment along with three other heuristic assignments (all the four of these assignments consume the same amount of power on average).

One can extend our preliminary results here by finding the best (or good heuristics) for linear networks with uniformly distributed nodes, and networks with nodes distributed according to a non-homogeneous Poisson process.

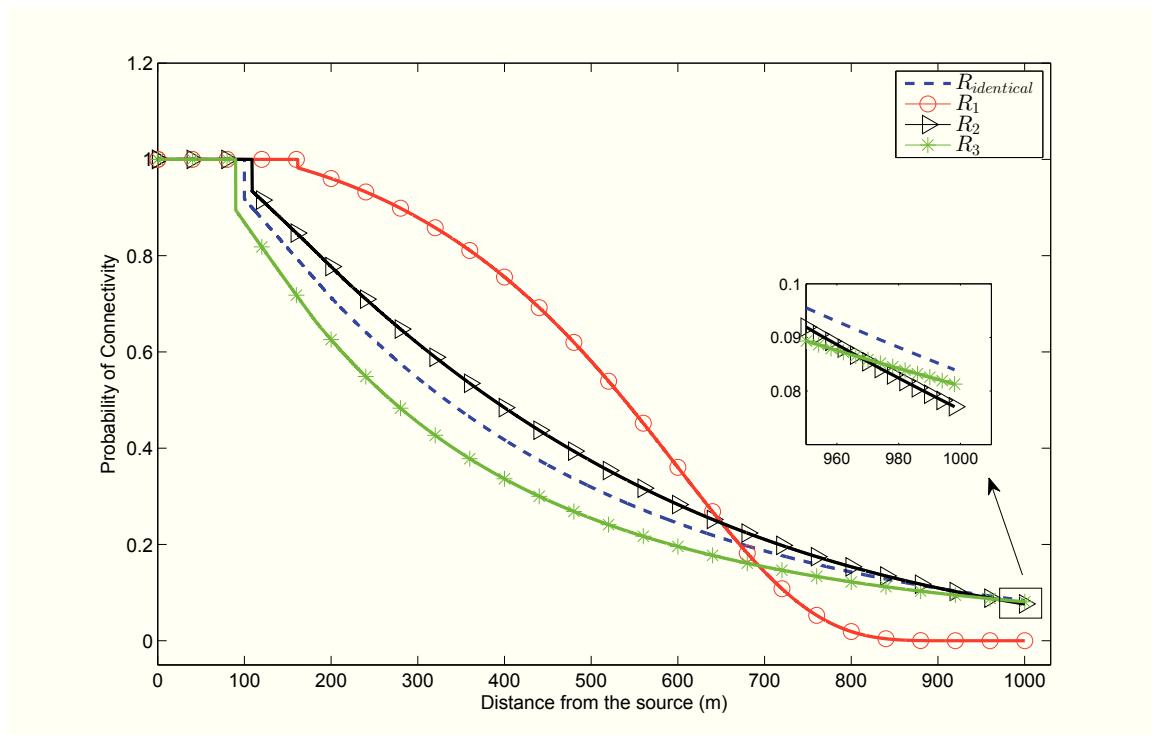


Figure 8.2: Probability of connectivity of a network with uniform node distribution employing the transmission range assignments shown in Fig. 8.1.

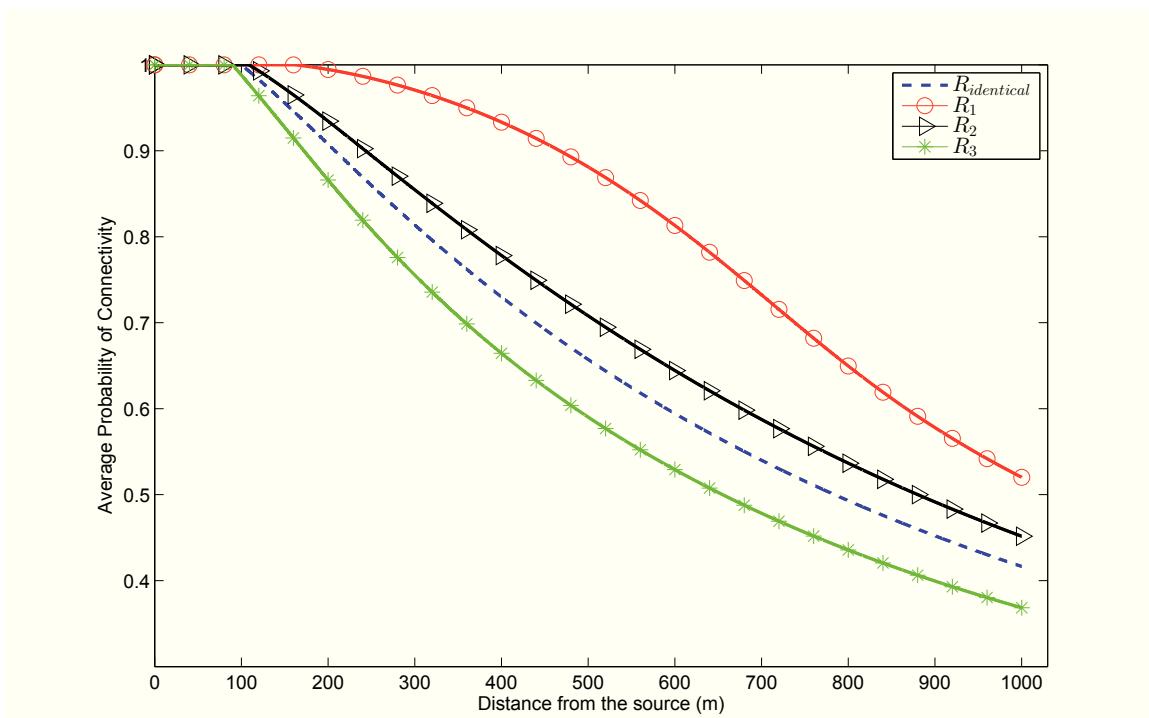


Figure 8.3: Average probability of connectivity up to the location of the x-axis in a network with uniform node distribution employing the transmission range assignments shown in Fig. 8.1.

List of References

- [1] A. Sharif-Nassab and F. Ashtiani, “Connectivity analysis of one-dimensional ad hoc networks with arbitrary spatial distribution for variable and fixed number of nodes,” *IEEE Transactions on Mobile Computing*, vol. 11, no. 10, pp. 1425–1435, Oct. 2012.
- [2] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [3] K. Pahlavan and A. Levesque, *Wireless Information Networks*. New York: Wiley-Interscience, 1995.
- [4] C. de Morais Cordeiro and D. P. Agrawal, *Ad hoc and sensor networks: theory and applications*. World Scientific, 2011.
- [5] E. Royer and T. Chai-Keong, “A review of current routing protocols for ad hoc mobile wireless networks,” *IEEE Personal Communications Magazine*, vol. 6, no. 2, pp. 46–55, 1999.
- [6] L. M. Kirousis, E. Kranakis, D. Krizanc, and A. Pelc, “Power consumption in packet radio networks,” *Theoretical Computer Science*, vol. 243, pp. 289–305, July 2000.
- [7] M. Xie and M. Haenggi, “Towards an end-to-end delay analysis of wireless multihop networks,” *Ad Hoc Networks*, vol. 7, no. 5, pp. 849–861, 2009.
- [8] Y. Wang, M. C. Vuran, and S. Goddard, “Cross-layer analysis of the end-to-end delay distribution in wireless sensor networks,” *IEEE/ACM Transactions on Networking*, vol. 20, no. 1, pp. 305–318, 2012.
- [9] M. Cagalj, J. Hubaux, and C. Enz, “Minimum-energy broadcast in all-wireless networks: Np-completeness and distribution issues,” in *Proc. of ACM MobiCom02*, p. 172182, Sept. 2002.
- [10] A. Boukerche, H. O. E. Nakamura, and A. Loureiro, “Localization systems for wireless sensor networks,” *IEEE Wireless Communications*, vol. 14, pp. 6–12, Dec. 2007.
- [11] G. Han, H. Xu, T. Q. Duong, J. Jiang, and T. Hara, “Localization algorithms of wireless sensor networks: a survey,” *Telecommunication Systems*, vol. 52, pp. 2419–2436, 2013.

- [12] J. Wang, R.K.Ghosh, and S. K. Das, “A survey on sensor localization,” *Journal of Control Theory and Applications*, vol. 8, pp. 2–11, 2010.
- [13] E. Ekici, S. Vural, J. McNair, and D. Al-Abri, “Secure probabilistic location verification in randomly deployed wireless sensor networks,” *Ad Hoc Networks*, vol. 6, pp. 195–209, 2008.
- [14] E. Ekici, J. McNair, and D. Al-Abri, “A probabilistic approach to location verification in wireless sensor networks,” in *Proc. of IEEE International Conference on Communications (ICC'06)*, vol. 8, pp. 3485–3490, 2006.
- [15] Y. Zeng, J. Cao, J. Hong, S. Zhang, and L. Xie, “Secure localization and location verification in wireless sensor networks: a survey,” *The Journal of Supercomputing*, vol. 64, no. 3, pp. 685–701, 2013.
- [16] C. A. Ardagna, M. Cremonini, E. Damiani, S. D. C. di Vimercati, and P. Samarati, “Supporting location-based conditions in access control policies,” in *Proceedings of the 2006 ACM Symposium on Information, Computer and Communications Security*, (New York, USA), pp. 212–222, 2006.
- [17] H. Hartenstein and K. P. Laberteaux, “A tutorial survey on vehicular ad hoc networks,” *IEEE Communications Magazine*, vol. 46, no. 6, pp. 164–171, June 2008.
- [18] I. Akyildiz, S. Weilian, Y. Sankarasubramaniam, and E. Cayirci, “A survey on sensor networks,” *IEEE Communications Magazine*, vol. 40, pp. 102–114, Aug 2002.
- [19] L. Xiao and F. Gao, “A comprehensive review of the development of adaptive cruise control (acc) systems,” *Vehicle System Dynamics*, vol. 48, pp. 1167–1192, April 2010.
- [20] E. Coelingh and S. Solyom, “All aboard the robotic road train,” *IEEE Spectrum*, vol. 49, pp. 34–39, November 2012.
- [21] A. Casteigts, A. Nayak, and I. Stojmenovic, “Communication protocols for vehicular ad hoc networks,” *Wireless Communications Mobile Computing*, vol. 11, pp. 567–582, May 2011.
- [22] Y. Yang, D. Divan, R. G. Harley, and T. G. Habetler, “Power line sensornet-a new concept for power grid monitoring,” in *IEEE Power Engineering Society General Meeting*, pp. 1–8, June 2006.
- [23] V. Gungor, L. Bin, and G. Hancke, “Opportunities and challenges of wireless sensor networks in smart grid,” *IEEE Transactions on Industrial Electronics*, vol. 57, pp. 3557–3564, October 2010.
- [24] A. E. F. Clementi, P. Crescenzi, P. Penna, G. Rossi, and P. Vocca, “On the complexity of computing minimum energy consumption broadcast subgraphs,” in *Proc. of the 18th Annual Symposium on Theoretical Aspects of Computer Science (STACS)*, pp. 121–131, 2001.

- [25] A. E. F. Clementi, P. Penna, and R. Silvestri, "On the power assignment problem in radio networks," *Mobile Networks and Applications*, vol. 2, pp. 125–140, April 2004.
- [26] S. Vural and E. Ekici, "Probability distribution of multi-hop-distance in one-dimensional sensor networks," *Computer Networks*, vol. 51, pp. 3727–3749, 2007.
- [27] A. Zanella, G. Pierobon, and S. Merlin, "On the limiting performance of broadcast algorithms over unidimensional ad-hoc radio networks," in *Proc. of WMPC*, pp. 165–169, 2004.
- [28] D. Niculescu and B. Nath, "Dv based positioning in ad hoc networks," *Telecommunication Systems*, vol. 22, no. 1-4, pp. 267–280, 2003.
- [29] G. Wu, S. Wang, B. Wang, Y. Dong, and S. Yan, "A novel range-free localization based on regulated neighborhood distance for wireless ad hoc and sensor networks," *Computer Networks*, vol. 56, no. 16, pp. 3581–3593, 2012.
- [30] B. Wang, G. Wu, S. Wang, and L. T. Yang, "Localization based on adaptive regulated neighborhood distance for wireless sensor networks with a general radio propagation model," *IEEE Sensors Journal*, vol. 14, pp. 3754–3762, Nov. 2014.
- [31] J. E. Wieselthier, G. D. Nguyen, and A. Ephremides, "On the construction of energy-efficient broadcast and multicast trees in wireless networks.,," in *Proc. of the 19th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, vol. 2, p. 585594, 2000.
- [32] S. Guo and O. Yang, "Energy-aware multicasting in wireless ad hoc networks: A survey and discussion," *Computer Communications*, vol. 30, pp. 2129–2148, 2007.
- [33] I. Caragiannis, C. Kaklamanis, and P. Kanellopoulos, "New results for energy-efficient broadcasting in wireless networks," in *Proc. of the 13th Annual International Symposium on Algorithms and Computation*, pp. 332–343, 2002.
- [34] A. E. F. Clementi, M. D. Ianni, and R. Silvestri, "The minimum broadcast range assignment problem on linear multi-hop wireless networks," *Theoretical Computer Science*, vol. 299, pp. 751–761, April 2003.
- [35] C. Ambuhl, A. E. F. Clementi, M. D. Ianni, G. Rossi, A. Monti, and R. Silvestri, "The range assignment problem in non-homogeneous static ad-hoc networks," in *Proc. of Parallel and Distributed Processing Symposium*, p. 224, April 2004.
- [36] J. E. Wieselthier, G. D. Nguyen, and A. Ephremides, "Energy-efficient broadcast and multicast trees in wireless networks," *Mobile Networks and Applications*, vol. 7, pp. 481–492, 2002.
- [37] P. J. Wan, G. Calinescu, X. Y. Li, and O. Frieder, "Minimum-energy broadcast routing in static ad hoc wireless networks," in *Proc. of 20th INFOCOM*, pp. 1162–1171, 2001.

- [38] C. Diot, W. Dabbous, and J. Crowcroft, "Multipoint communication: a survey of protocols, functions, and mechanisms," *IEEE Journal on Selected Areas in Communications*, vol. 15, pp. 277–290, Apr 1997.
- [39] J. Bauer, D. Haugland, and D. Yuan, "New results on the time complexity and approximation ratio of the broadcast incremental power algorithm," *Information Processing Letters*, vol. 109, pp. 615–619, 2009.
- [40] D. Yuan, J. Bauer, and D. Haugland, "Minimum-energy broadcast and multicast in wireless networks: An integer programming approach and improved heuristic algorithms," *Ad Hoc Networks*, vol. 6, pp. 696–717, 2008.
- [41] G. D. Nguyen, "General algorithms for construction of broadcast and multicast trees with applications to wireless networks," *Journal of Communications and Networks*, vol. 7, pp. 263–277, Sept. 2005.
- [42] A. Murata and A. Matsubayashi, "Minimum energy broadcast on rectangular grid wireless networks," *Theoretical Computer Science*, vol. 412, pp. 5167–5175, 2011.
- [43] T. Calamoneri, A. Clementi, M. D. Ianni, M. Lauria, A. Monti, and R. Silvestri, "Minimum-energy broadcast and disk cover in grid wireless networks," *Theoretical Computer Science*, vol. 399, p. 3853, 2008.
- [44] G. K. Das, S. C. Ghosh, and S. C. Nandy, "Improved algorithm for minimum cost range assignment problem for linear radio networks," in *Proc. of International Workshop on Distributed Computing (IWDC 2004)*, pp. 412–423, 2004.
- [45] P. Piret, "On the connectivity of radio networks," *IEEE Transactions on Information Theory*, vol. 37, pp. 1490–1492, Sep 1991.
- [46] P. Santi and D. M. Blough, "The critical transmitting range for connectivity in sparse wireless ad hoc networks," *IEEE Transactions on Mobile Computing*, vol. 2, no. 1, pp. 25–39, 2003.
- [47] M. Artimy, "Local density estimation and dynamic transmission-range assignment in vehicular ad hoc networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 8, no. 3, pp. 400–412, Sept. 2007.
- [48] V. K. M. Ajeer, P. C. Neelakantan, and A. V. . Babu, "Network connectivity of one-dimensional vehicular ad hoc network," in *Proc. of Communications and Signal Processing (ICCPSP)*, pp. 241 – 245, Feb. 2011.
- [49] A. Ghasemi and S. Nader-Esfahani, "Exact probability of connectivity in one-dimensional ad hoc wireless networks," *IEEE Communications Letters*, vol. 10, no. 4, pp. 251–253, April 2006.
- [50] D. Miorandi and E. Altman, "Connectivity in one-dimensional ad hoc networks: a queueing theoretical approach," *Wireless Networks*, vol. 12, no. 5, pp. 573–587, 2006.

- [51] S. Yousefi, E. Altman, R. El-Azouzi, and M. Fathy, "Analytical model for connectivity in vehicular ad hoc networks," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 6, pp. 3341–3356, 2008.
- [52] H. I.W.-H, K. Leung, and J. W. Polak, "Stochastic model and connectivity dynamics for vanets in signalized road systems," *IEEE/ACM Transactions on Networking*, vol. 19, pp. 195–208, Feb. 2011.
- [53] H. I.W.-H, K. Leung, and J. W. Polak, "Connectivity dynamics for vehicular ad-hoc networks in signalized road systems," in *Teletraffic Congress*, pp. 1–8, Sept. 2009.
- [54] G. Han and A. Makowski, "One-dimensional geometric random graphs with non-vanishing densities -part i: A strong zero-one law for connectivity," *IEEE Transactions on Information Theory*, vol. 55, pp. 5832–5839, Dec. 2009.
- [55] G. Han and A. M. Makowski, "On the critical communication range under node placement with vanishing densities," in *Proc. of IEEE Int. Symp. on Information Theory (ISIT 2007)*, pp. 831–835, Jun. 2007.
- [56] S. Vural and E. Ekici, "Analysis of hop-distance relationship in spatially random sensor networks," in *Proceedings of the 6th ACM international symposium on Mobile ad hoc networking and computing*, pp. 320–331, 2005.
- [57] B. Hofmann-Wellenhof, H. Lichtenegger, and J. Collins, *Global Positioning System: Theory and Practice*. Springer-Verlag, 1997.
- [58] T. Kunz and B. Tatham, "Localization in wireless sensor networks and anchor placement," *Journal of Sensor and Actuator Networks*, vol. 1, pp. 36–58, 2012.
- [59] M. Soleimanifar, X. Shen, M. Lu, and I. Nikolaidis, "Applying received signal strength based methods for indoor positioning and tracking in construction applications," *Canadian Journal of Civil Engineering*, vol. 41, no. ja, pp. 703–716, 2014.
- [60] Y. Wang, F. Zheng, M. Wiemeler, W. Xiong, and T. Kaiser, "Reference selection for hybrid toa/rss linear least squares localization," in *Proc. of Vehicular Technology Conference (VTC Fall)*, pp. 1–5, IEEE, 2013.
- [61] A. F. Molisch, D. Cassioli, C.-C. Chong, S. Emami, A. Fort, B. Kannan, J. Karedal, J. Kunisch, H. G. Schantz, K. Siwiak, *et al.*, "A comprehensive standardized model for ultrawideband propagation channels," *IEEE Transactions on Antennas and Propagation*, vol. 54, no. 11, pp. 3151–3166, 2006.
- [62] S. Marano, W. M. Gifford, H. Wymeersch, and M. Z. Win, "Nlos identification and mitigation for localization based on uwb experimental data," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 7, pp. 1026–1035, 2010.
- [63] P. Bahl and V. Padmanabhan, "Radar: An in-building rf-based user location and tracking system," in *Proc. of 19th INFOCOM*, 2000.

- [64] L. Li and T. Kunz, "Cooperative node localization using nonlinear data projection," *ACM Transactions on Sensor Networks (TOSN)*, vol. 5, no. 1, pp. 1–26, 2009.
- [65] V. Ramadurai and M. L. Sichitiu, "Localization in wireless sensor networks: A probabilistic approach," in *Proc. of ICWN 2003*, pp. 275–281, June 2003.
- [66] R. Peng and M. L. Sichitiu, "Probabilistic localization for outdoor wireless sensor networks," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 11, pp. 53–64, 2007.
- [67] M. R. Ataei, A. H. Banihashemi, and T. Kunz, "Low-complexity energy-efficient broadcasting in one-dimensional wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 61, pp. 3276–3282, Sept 2012.
- [68] D. B. West, *Introduction to Graph Theory*. Prentice Hall, 2001.
- [69] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*. The MIT press, July 2009.
- [70] M. R. Ataei, A. H. Banihashemi, and T. Kunz, "Minimum-energy broadcasting for cross wireless ad-hoc networks," in *Proceedings of IEEE ICC 2015*, June 2015.
- [71] V. F. Kolchin, B. A. Sevastyanov, and V. P. Chistyakov, *Random allocations*. V.H. Winston and Sons, 1978.
- [72] M. R. Ataei, A. H. Banihashemi, and T. Kunz, "An energy-efficient transmission range assignment for 1-d wireless ad hoc networks," *IEEE Wireless Communications Letters*, vol. 2, no. 5, pp. 543–546, Oct. 2013.
- [73] S. M. Ross, *Introduction to Probability Models*. Academic Press, 10th ed., 2009.
- [74] R. Misra and C. Manda, "Minimum connected dominating set using a collaborative cover heuristic for ad hoc sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 21, no. 3, pp. 292–302, March 2010.
- [75] K. G. Binmore, *Mathematical Analysis: A Straightforward Approach*. Cambridge University Press, 2 ed., Feb. 1983.
- [76] M. R. Ataei, T. Kunz, and A. H. Banihashemi, "Localization and location verification in non-homogeneous one-dimensional wireless ad-hoc networks," *IEEE Journal on Selected Areas in Communications, Special Issue on Location-Awareness for Radios and Networks*, vol. 33, no. 7, May 2015.
- [77] M. R. Ataei, T. Kunz, and A. H. Banihashemi, "Localization in non-homogeneous one-dimensional wireless ad-hoc networks," in *Proceedings of IEEE ICC 2015*, June 2015.
- [78] P. Vigneron and J. Pugh, "Propagation models for mobile terrestrial vhf communications," in *IEEE Military Communications Conference (MILCOM)*, pp. 1–7, IEEE, 2008.
- [79] H. L. Van Trees and K. L. Bell, *Detection Estimation and Modulation Theory*. John Wiley & Sons, 2013.