REASONING ABOUT WIRELESS PROTOCOL BEHAVIOR

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of the Ohio State University

By

Tae Woo Kwon, BS/MS
Graduate Program in Computer Science and Engineering

The Ohio State University
2012

Dissertation Committee:

Anish Arora, Advisor
Emre Ertin
Emre Koksal
Kannan Srinivasan
© Copyright by
Tae Woo Kwon
2012
ABSTRACT

The capability of sensor nodes has been improved tremendously in the last decade. However, wireless sensor network (WSN) protocol performance still not considered reliable, predictable, or even repeatable. Although performance repeatability is a fundamental requirement in the network protocol development, in many cases, wireless protocols, which are extensively tested and well tuned in the WSN testbeds, misbehave in the actual target environments. In this dissertation, we study how to achieve performance repeatability of WSN protocols. There are a couple of crucial reasons why WSN protocol performance is not considered repeatable: unreliable hardware, software, and low power radio communication. Moreover, network protocol behaviors can themselves exhibit nontrivial variability, and this variability may only be inadequately understood in the testing phase. The multi-faceted difficulty with ensuring desired protocol behavior in the field coupled with the high cost of testing and tuning the performance in the field, motivates the scientific study of tools and techniques for reproducing network behavior across test and deployment environments.

To attack this performance unrepeatability problem of WSN protocols, we approach solutions with analytical and data driven methods. With analytical method, we characterize the wireless protocol performance and behavior mathematically. We try to identify uncertainty factors of WSN environments such that WSN testbeds and deployments. In the next step, we analyze their impact on protocol performance and behavior. Towards achieving reproducible performance across networks
of potentially different environments, we adopt the concept of realizing the same (or measurably close to the same) “link usage spectrum”, defined as the probability distribution with which the network protocol selects links of different length from among all the available links in the network at hand. Based on the mathematical modeling of link usage spectrum, we derive a closed form equation of the expected performance and the variance of wireless protocols using $PRR \times D$ for routing metric. Equipped with mathematical modeling of link usage spectrum and the expected performance of wireless protocols, we provide methods to reproduce the comparable protocol performance across environments by matching link usage spectrums as close as possible (method 1) or by matching the expected performances as close as possible (method 2) of two different environments. These two methods work well with 1-dimensional chain topologies and 2-dimensional grid topology, but are not applicable to topologies except chain or grid. This problem can be solved with data driven method.

With data driven method, we simulate wireless protocols with operational models with detailed link quality data collected with a WSN testbed resource specification profiling program, RS-Profiler. It is difficult to analytically model effects such as multi-path and component variability. However, models based on measurement data that captures these effects can improve accuracy substantially. We implement RS-Profiler that collects RF data of all links and all nodes, e.g. RSSI, SNR, PRR, noise floor, efficiently. We provide two performance prediction algorithms (operational models of routing protocols) that accurately predict the expected performance of protocols using the cumulative routing metric (e.g. ETX) and the 1-hop routing metric (e.g. $PRR \times D$) based on the collected RF resource specification. These two algorithms can be applied to any topologies. We prove that performance repeatability within WSN testbed and also across WSN testbed with performance prediction algorithms and RF resource specification with extensive experiments over 18 different
2-dimensional grid networks. However these data driven methods incur scalability problems in data collection. We present a study on time complexity of RS-Profiling, which is $O(N)$ with the number of nodes $N$. Because the profiling time will be impractically long (e.g. 3 days for full RSSI, noise floor, and PRR for all 16 channels, for all 8 different transmission power levels, for all links of all nodes), we present studies of three methods to relieve this RS-Profiling time scalability problem.

Performance repeatability of wireless protocols across environments can be achieved not only by matching the expected performance, but also by achieving the comparable performance variances. Therefore, we study methods to predict the variance of WSN protocols with mathematical equation and with Monte-Carlo simulation. We also study the upper/lower bounds of protocol performance in a given communication environment. We validate our protocol variance study with experimental results.
To god who guides me...

to my late father who had burning thirst of study...

to my wife, twin sons...

to my mother, father in law, mother in law...

for their endless love, support and patience.
ACKNOWLEDGMENTS

My Ph.D program in Ohio State University has not been easy, and would have been fruitless without the tremendous help from my advisor, Prof. Anish Arora, and my co-advisor, Prof. Emre Ertin. Prof. Arora always guides me with wisdom and through understanding on computer system theory that I could get good solutions for really challenging and hard research questions. Prof. Arora also helped and advised me to write academic papers and ultimately publish them in highly valued conferences and journals. Prof. Ertin gave me a crucial idea to derive the fundamental theorems of my Ph.D thesis. And, he always cheers me to get through hard times that I suffered during my Ph.D program. Prof. Ertin supported me to attend and present a paper in Adhocnets’2010, and ultimately the upgraded paper could be published through Elsevier Journal. I also want to give thanks to my colleagues in my lab who always give me helps and kind replies on my frequent questions and askings.

Taewoo Kwon

May, 2012
VITA

Born in Taejeon, Republic of Korea

BS in Computer Science, Yonsei University, Seoul, Korea

MS in Electrical and Computer Engineering, University of Massachusetts, MA, USA

Graduate Teaching / Research Associate, The Ohio State University

PUBLICATIONS

Journal Papers:


Peer Reviewed Conferences:


Posters & Short Papers


Technical Reports

FIELDS OF STUDY

Major Field: Computer Science and Engineering

Specialization: Computer Networks and Systems
# TABLE OF CONTENTS

Abstract ................................................................. ii
Dedication ............................................................... iv
Acknowledgments ......................................................... vi
Vita ....................................................................... vii
List of Figures ........................................................... xiii

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction ............................................. 1</td>
</tr>
<tr>
<td>1.1</td>
<td>Research questions ......................................... 1</td>
</tr>
<tr>
<td>1.2</td>
<td>Challenges .................................................. 2</td>
</tr>
<tr>
<td>1.3</td>
<td>Previous work related to performance repeatability .......... 3</td>
</tr>
<tr>
<td>1.3.1</td>
<td>Characterizing the protocol performance .................... 3</td>
</tr>
<tr>
<td>1.3.2</td>
<td>Characterizing the link selection behavior .................. 4</td>
</tr>
<tr>
<td>1.3.3</td>
<td>Performance repeatability of WSN protocols ................ 4</td>
</tr>
<tr>
<td>1.4</td>
<td>Our contributions .......................................... 6</td>
</tr>
<tr>
<td>2</td>
<td>Link Usage Spectrum and Consistent Reproduction of Wireless Routing Protocol Performance .... 12</td>
</tr>
<tr>
<td>2.1</td>
<td>Link usage spectrum and network transplant error ............. 13</td>
</tr>
<tr>
<td>2.1.1</td>
<td>Comparing protocol performance in two settings ............ 16</td>
</tr>
<tr>
<td>2.2</td>
<td>Experimental study of protocol transplantation ............... 18</td>
</tr>
<tr>
<td>2.2.1</td>
<td>Experimental setup ......................................... 18</td>
</tr>
<tr>
<td>2.2.2</td>
<td>Physical &amp; link layer ....................................... 19</td>
</tr>
<tr>
<td>2.2.3</td>
<td>Messaging layer ........................................... 21</td>
</tr>
<tr>
<td>2.2.4</td>
<td>Results ................................................... 21</td>
</tr>
<tr>
<td>2.3</td>
<td>Analytical methods for predicting link usage spectrum .......... 23</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Analytical model for one-dimensional and two-dimensional uniform graphs ..................... 24</td>
</tr>
<tr>
<td>2.3.2</td>
<td>Validation of analytical approximations ...................... 28</td>
</tr>
</tbody>
</table>

3.1 Wireless protocol performance modeling
   3.1.1 Calculating the expected performance metrics
   3.1.2 Calculating the moments of the conditional SNR
   3.1.3 Tuning power for matching performance
   3.1.4 Calculating the variance of performance metrics
   3.1.5 Upper/lower bounds of the Protocol performance

3.2 Validation
   3.2.1 Micro study: chain one-dimensional topology
   3.2.2 Macro study: grid two-dimensional topology

3.3 Conclusion

4 Performance Repeatability of Wireless Protocols with Data Driven Method using RF Resource Specification

4.1 Resource specification data collection with Profiler
   4.1.1 RS-Profiler details
   4.1.2 Scalability problem of RS-Profiler

4.2 Accurate WSN routing protocol performance prediction based on RF resource specification
   4.2.1 Performance prediction for 1-hop metric
   4.2.2 Performance prediction for cumulative metric
   4.2.3 Validation

4.3 Performance repeatability of WSN routing protocols in WSN testbeds
4.4 Temporal Link Quality Variance Study
4.5 Predicting performance variance with resource specification
   4.5.1 Monte-Carlo simulation study with performance prediction
   4.5.2 Validation of performance variance prediction methods
   4.5.3 Validation of protocol variance prediction

4.6 Conclusion

5 Scalability of Resource Specification Profiling on Large Scale WSN Testbeds

5.1 Complexity analysis of resource specification profiling time
5.2 Accurate prediction of RSSI values on different transmission power levels
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>14</td>
</tr>
<tr>
<td>2.2</td>
<td>20</td>
</tr>
<tr>
<td>2.3</td>
<td>22</td>
</tr>
<tr>
<td>2.4</td>
<td>29</td>
</tr>
<tr>
<td>2.5</td>
<td>30</td>
</tr>
<tr>
<td>2.6</td>
<td>30</td>
</tr>
<tr>
<td>2.7</td>
<td>31</td>
</tr>
<tr>
<td>2.8</td>
<td>33</td>
</tr>
<tr>
<td>2.9</td>
<td>35</td>
</tr>
<tr>
<td>2.10</td>
<td>37</td>
</tr>
<tr>
<td>2.11</td>
<td>37</td>
</tr>
<tr>
<td>2.12</td>
<td>39</td>
</tr>
<tr>
<td>3.1</td>
<td>54</td>
</tr>
</tbody>
</table>
3.2 SNR of the chosen/all links. Analytical (Theorem 4). C:Chosen, A:All 55
3.3 PRR of the chosen/all links. Analytical (Theorem 3) . . . . . . . . . 56
3.4 Cumulative link usage spectrum. (1) analytical (Theorem 2), (2) ex-
periment . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 56
3.5 (1) Average number of transmissions (experimental), (2) average link
length (experimental) . . . . . . . . . . . . . . . . . . . . . . . . . . . . 57
3.6 Cumulative usage weights at KanseiGenie. (1) analytical, (2) exper-
imental. transmission power = (+:0, ×:-5, □:-10, ■:-15) dBm . . . . . . 60
3.7 Cumulative usage weights at Motelab. (1) analytical, (2) experimental. transmission power = (+:0, ×:-5, □:-10, ■:-15) dBm . . . . . . 60
3.8 Cumulative usage weights at Tutornet. (1) analytical, (2) experimen-
tal. transmission power = (+:-5, ×:-7, □:-10, ■:-15) dBm . . . . . . 61
3.9 Cumulative usage weights at TWIST. (1) analytical, (2) experimental.
transmission power = (+:-5, ×:-10, □:-15, ■:-25) dBm . . . . . . . . . 61
3.10 Number of transmissions. (1) analytical, (2) experimental. K: Kan-
seiGenie, M: Motelab, TU: Tutornet, TW:TWIST . . . . . . . . . . . . 62
4.1 SNR vs. PRR scatter plot in KanseiGenie . . . . . . . . . . . . . . . 71
4.2 Analytical SNR vs. PRR graph and experimentally adjusted SNR vs.
PRR graph of KanseiGenie . . . . . . . . . . . . . . . . . . . . . . . . . 71
4.3 Measured SNR values on all pairs links of a 4 × 4 grid topology with
channel 11, transmission power -15 dBm in KanseiGenie . . . . . . . 72
4.4 Expected PRR values for all pairs links of a 4 × 4 grid topology with
channel 11, transmission power -15 dBm in KanseiGenie . . . . . . . 73
4.5 Measured SNR values on all pairs links of a 4 × 4 grid topology with
channel 11, transmission power -25 dBm in KanseiGenie . . . . . . . 73
4.6 Expected PRRs for all pairs links of a 4 × 4 grid topology with channel
11, transmission power -25 dBm in KanseiGenie . . . . . . . . . . . . 74
4.7 Expected $P R R \times D$ values at source (node1) with channel 11, trans-
mission power -15 dBm . . . . . . . . . . . . . . . . . . . . . . . . . 74
4.8 Expected $P R R \times D$ values at the source (node1) with channel 11,
transmission power -25 dBm . . . . . . . . . . . . . . . . . . . . . . . . 75

xiv
4.9 Expected $P_{RR} \times D$ values at node 4 with channel 11, transmission power -25 dBm ........................................ 76

4.10 Expected ETX values for all pairs links of a 4 × 4 grid topology with channel 11, transmission power -15 dBm in KanseiGenie ............ 77

4.11 Expected ETX values at the source (node1) with channel 11, transmission power -15 dBm ........................................ 77

4.12 Expected ETX values for all pairs links of a 4 × 4 grid topology with channel 11, transmission power -25 dBm in KanseiGenie ............ 79

4.13 Expected ETX values at the source (node1) with channel 11, transmission power -25 dBm ........................................ 79

4.14 Long term temporal SNR variance of links of Candidate 8 (1), Candidate 10 (2) at KanseiGenie channel 11 ......................... 83

4.15 Short term temporal RSSI variance of links of Candidate 8 (1), Candidate 10 (2) at KanseiGenie channel 11 ......................... 85

4.16 Path loss at KanseiGenie in channel 11 ........................................ 88

5.1 Cumulative prediction errors based on raw RSSI and calibrated RSSI: (1) 0 dB vs. -1 dB, (2) 0 dB vs. -3 dB, (3) 0 dB vs. -5 dB, (4) 0 dB vs. -7 dB, (5) 0 dB vs. -10 dB, (6) 0 dB vs. -15 dB, (7) 0 dB vs. -25 dB .... 96

5.2 Prediction errors based on calibrated RSSI: (1) 0 dB vs. -1 dB, (2) 0 dB vs. -3 dB, (3) 0 dB vs. -5 dB, (4) 0 dB vs. -7 dB, (5) 0 dB vs. -10 dB, (6) 0 dB vs. -15 dB, (7) 0 dB vs. -25 dB ..................... 97

5.3 Noise floor and interference measurements at a node of KanseiGenie . 99

5.4 Noise floor of channel 26, 11, 12 at KanseiGenie ..................... 100

5.5 RSSI vs. distance: KanseiGenie channel 26 (1) and Channel 12 (2) . 101

5.6 RSSI, noise floor and interference measurements of link1 at a node of KanseiGenie in channel 23 ........................................ 102

5.7 RSSI, noise floor and interference measurements of link3 at a node of KanseiGenie in channel 23 ........................................ 103

5.8 RSSI Measurements of links at a node of KanseiGenie channel 23 .. 104

5.9 Cumulative RSSI prediction errors at a node of KanseiGenie in channel 23. (1) predicted with static link qualities, (2) predicted with average noise floor difference ........................................ 104
5.10 Cumulative PRR prediction errors with links of a 4×4 grid network of KanseiGenie in channel 11 with transmission power of -25 dBm . . 106

5.11 SNR vs. PRR prediction error with links of a 4×4 grid network of KanseiGenie in channel 11 with transmission power of -25 dBm . . . 106
CHAPTER 1
INTRODUCTION

More than a decade has passed since the emergence of “Sensor Network” system. The capability of sensors has improved to be comparable to the early stage personal computers. Now, sensor nodes are powerful enough to support virtual machines [23] [10], and multi-threaded [19] operation. However, wireless sensor network (WSN) protocol performance is still not considered reliable, predictable, or even repeatable.

1.1 Research questions

Performance repeatability is a fundamental and general requirement of the network protocol development. However, experiences in deploying wireless sensor networks during this decade have yielded a number of surprises, wherein network behavior in the field diverged substantially from that seen in laboratory tests.

With many obstacles, to achieve performance repeatability of WSN protocols, we need a systematic and scientific method to interpret the experimental result and to predict the expected result of wireless protocols in the target environment to reduce cost and time to do experiment repeatedly. The systematic and scientific study is reasoning about wireless protocol behavior to answer questions:

- What are the factors impacting performance metric?
- What is the mean and variance of a WSN performance metric?
• Given a target network and a test network which are identical modulo of their values for some factors, can the test network be controlled to have, with high probability, the same metric value as the target network?

• How can we generalize and automate the process achieving the performance repeatability within and across environments?

1.2 Challenges

There are several reasons that WSN protocol performance is considered hard to reproduce. First, the hardware and software system of sensor network system are unreliable. About 30% of sensor nodes manufactured from factory are bad receivers, which perform poor in receiving packets. Software system is also unreliable. Due to the limited capacity of memory and CPU power of sensor nodes and also the limited energy source, sensor software has been designed light-weight, and lack of the fault tolerance mechanisms and the troubleshooting mechanisms. Because, debugging is extremely hard (there is no good debugger for sensors and sometimes we resort on 3 LED lights for debugging) for sensor protocol development, it is very hard to root-cause the bugs and faults.

Low-power communication also contributes to performance unrepeatability of WSN protocols. WSN communication is notorious for unreliable and complex. Due to the multi-path effect, especially indoor, with the same node sets and the same transmission power, and with the same communication distance, the link qualities (e.g. RSSI, SNR, and PRR) vary highly according to the location. And, due to the ambient noise variance and the interference variance, the link qualities vary with time even with the same node sets and the same transmission powers and distance.

On top of those hardware, software and the radio communication factors, we
also have to consider the behavior of the specific protocols according to the given environments. For example, the performance of routing protocols, which choose the best links according to their routing metric, vary according to the communication environments.

To achieve the performance repeatability of WSN protocols, we have to characterize the behavior and the performance of WSN protocols according to the communication environments, and also find the methods to measure the important RF data of communication environments efficiently.

1.3 Previous work related to performance repeatability

1.3.1 Characterizing the protocol performance

For wired protocols, there have been many studies on their performance prediction. Golestani [16], [17], Parekh [30], Cobb [14] studied fundamental methods to characterize the wired protocol performances, which enabled reasoning about the wired protocol behaviors. However, because the communication physics of wireless protocols is fundamentally different from that of wired protocols, the previous studies on wired protocol behavior cannot be applied to wireless protocols.

Chiasserini et al. [12] proposed Markov model of a protocol performance of wireless protocols in a duty-cycled network, but they assumed static routing, in other words fixed routing, which is not the general option for WSN protocols. Zuniga et al. [38] calculated the expected performance of a routing protocol that choose paths based on the \( PRR \times D \) metric, but they could not provide a closed form equation to calculate the expected performance and the probability that a link will be chosen over all other links. They provided the results from numerical simulations instead.
1.3.2 Characterizing the link selection behavior

There have been a few previous works on the link selection behavior. There is significant literature however, e.g. [37], that models the bit error rate for radio channels and thus calculates performance metrics such as signal to noise ratio (SNR), packet reliability rate (PRR), expected number of transmissions (ETX), $\text{PRR} \times D$ (the forwarding distance) [31], and the expected latency per unit distance ($\text{ELD} = \frac{1}{\text{PRR} \times D}$) [35].

There is a related work that implicitly exploits the link usage spectrum idea [31], although its role is different: the spectrum is used as a tool for calculating average network metrics that are in turn used for choosing between protocols and optimizing a protocol realization with respect to its intended forwarding metric.

[31] quantified the difference between their local optimum metric, $\text{PRR} \times D$ and the global optimal metric ETX ([15]). [31] observed that there is only narrow difference between ETX and $\text{PRR} \times d$ on the energy performance ($\approx \frac{r}{t}$), where $r$ is the delivery rate and $t$ is the total number of transmission, for different densities and network sizes. In calculating average network performance metric ($E[\xi_{\text{PRR} \times D}]$), equation (25) of [38] appear to use incorrectly the unconditional mean PRR of a link of length $d$ ($E[\text{PRR}|d]$), instead of the conditional mean $E[\text{PRR}|d_f = d]$ given the event ($d_f = d$) that the link of length $d$ was chosen by the forwarding protocol over all other link lengths. In general $E[\text{PRR}|d_f = d] > E[\text{PRR}|d]$ and accurate calculation of the network-wide performance metrics requires the use of conditional link qualities.

1.3.3 Performance repeatability of WSN protocols

Naik et al. [29] studied the spatial scaling problem within the same environment. They studied the performance repeatability into the different spatial scale, i.e. different inter-node distance, in the same environment. They compared the performance metrics of two WSN applications in two different scales in Kansei ([1]), an indoor
WSN testbed, by using different transmission power levels. In their experiments, Sprinkler [28], a protocol that provides a bulk data transmission service, and LOF [35], a protocol that provides a beacon-free routing service, were run on the testbed and then re-run in the same topology except that the distance between nodes was doubled. Sprinkler showed a similar performance with the metrics of the number of transmissions and latency, in both topologies, whereas LOF showed a non-negligible difference in performance with the metrics of per-hop link length, per-hop MAC latency, and end-to-end MAC latency. [29] concluded that across scaling within the same environment, an application favoring inner-band links will behave comparably while an application favoring middle-band links will behave differently; and across environments, all protocols are likely to behave differently. This negative conclusion clearly impacts the usefulness of testbeds as predictors of the protocol performance during the deployment across different environments.

Designing protocols with robust performance across different environments is a related research thrust. For example, Lin et al. [24] proposed a new link metric called competence capturing long term variation of link quality and designed a distributed route maintenance framework based on feedback control systems. [24] tried to achieve the performance repeatability by achieving a stable performance through choosing more stable links in long term, and equipping extra schemes (e.g. Retransmission, Transmission Power Control) to guarantee certain level of link reliability.

Our goal in this dissertation is to repeat performance of WSN protocols that may not select links with high competence.

Logging the detailed environmental information and using it for the performance repetition is another approach, and our work in this dissertation also follows this strategy. Luo et al. [26] designed and implemented EnviroLog, a distributed service networks via asynchronous event recording and replay. [26] argued that EnviroLog
improves repeatability of experimental testing of sensor network. However, EnviroLog can only repeat events except communication, therefore EnviroLog cannot achieve the repeatability of the protocol performance of wireless protocols, which varies on RF propagation environment.

1.4 Our contributions

To attack this performance unrepeatability problem of WSN protocols, we approach to solutions with analytical method and data driven method.

Analytical method. With analytical method, we characterize the wireless protocol performance and behavior mathematically. We try to identify uncertainty factors of WSN environments such that WSN Testbeds and deployments. Then, we analyze their impacts on protocol performance and behavior. One key uncertainty factor is that the effective topology of the laboratory tests is different from that of the field deployment: Not only the inter-node spatial (separation) scale in the two networks is different, but the environment signal propagation characteristics also tend to be different, and as a result the link selections and the intra-node traffic interference diverge. Differences in externally induced communication interference are another factor. Other scale differences in the field deployment, i.e., increasing the number of the nodes fielded, and consequent phase transition or instability issues are yet another factor. Moreover, network protocol behaviors can themselves exhibit non-trivial variability, and this variability may only be inadequately understood in the testing phase. The multi-faceted difficulty with ensuring desired protocol behavior in the field, coupled with the high cost of testing and tuning the performance in the field, motivates the scientific study of tools and techniques for reproducing network behavior across test and deployment environment. What do we mean by reproducing performance? Even if the test and deployment environments are the same and we
only displace the network in space, the realized network performance will be not identical. One can only achieve a probabilistic equivalence between two such networks. By probabilistic equivalence, we informally mean that the set of links exercised by the two networks are sampled from the same probability distribution. For two networks at different spatial scales in the same environment, positive results for achieving probabilistic equivalence have been presented earlier, i.e., by using transmission power control [29]. Power control was realized via attenuator hardware and/or software control. [29] also experimentally compared the performance metrics of two wireless sensor network (WSN) protocols — Sprinkler and LOF—at different spatial scales in KanseiGenie, an indoor WSN testbed [9] [2], to illustrate how to select the transmission power to achieve probabilistic equivalent behavior when scaling all inter-node distances by some constant. In contrast, for two networks at different spatial scales in different environments—in particular, with different path loss exponents—it is straightforward to show that it is impossible to achieve probabilistic equivalence using only transmission power control. This necessitates consideration of alternative techniques. Towards achieving comparable performance in networks in potentially different environments, we adopt the concept of realizing the same (or measurably close to the same) “link usage spectrum” in the networks. Informally speaking, the link usage spectrum of a network is the probability distribution with which the network protocol selects links of different length from among all the available links in the network at hand. There are two major factors that affect the link usage spectrum: the metric chosen by the network protocol for selecting parents to whom nodes forward their traffic and, more generically, the signal to noise ratio (SNR), or RSSI values of links. It is often the case that the chosen metric itself involves a function of the SNR (or RSSI) as well as the distance traversed by the link. In these cases, the link usage spectrum can be reformulated as a function of relative preference based on
SNR and the forwarding distance. We emphasize that this is only one of the ways to formulate the link usage spectrum, and that the analysis we perform subsequently in this dissertation is readily adapted to several other routing metrics.

Based on the mathematical modeling of link usage spectrum, we derive a closed form equation of the expected performance and the variance (chapter 3) of wireless protocols using $PRR \times D$ for routing metric. Our hypotheses on formulating the link usage spectrum and the expected performance are as follows. First, the received signal strength follows the log-normal fading and the normal distribution with the average values. Second, we assume topologies are infinite, so that we can apply the same equation to calculate the link usage spectrum and the expected performance toward destination for every node. Because of this assumption, our analytical results inevitably suffer the edge effect for nodes in boundary. Third, the background noise is relatively stable.

Equipped with mathematical modeling of link usage spectrum and the expected performance of wireless protocols, we provide methods to reproduce the comparable protocol performance across environments by matching link usage spectrums as close as possible (method 1) or by matching the expected performances as close as possible (method 2) of two different environments. The hypothesis of method 1 is that the link usage spectrum is a gross predictor of the performance of (a rich class of) network protocols. With this hypothesis, a network protocol will perform comparably in two network settings if the respective link usage spectrums of the protocol match closely in these settings. Said another way, even if the “available” link spectrum in the settings is different but the probability distribution of the chosen links is comparable, the protocol behavior in the settings will be comparable. The link usage spectrum can therefore be used to achieve predictable network behavior in the deployment setting, as follows. Consider the case where a prototype network is tested somewhere, say in an
indoor environment, before it is fielded elsewhere, say in an outdoor environment, with potentially different inter-node spacing. Since the indoor environment is persistent and easily instrumented for tests, it is relatively easy to collect fine-grain, long running protocol behavior information in it. Thus, if one could easily calculate the link usage spectrum data for the outdoor environment (by analysis, simulation, or experiment), then one could (analytically or experimentally) design the link usage spectrum for the indoor tests (say by choosing the indoor network spatial scaling facto and transmission power level) to be close to that of the field network. The resulting observed indoor protocol behavior would be predictive of the behavior to be observed in the field. We also provide experimental validation study for method 2 (performance repetition) with 2-dimensional $4 \times 4$ grid networks across four major WSN testbeds (KanseiGenie [2], Motelab [3], Tutornet [4], TWIST [5]).

**Data driven method.** With data driven method, we simulate wireless protocols with operational models with detailed link quality data collected with a WSN testbed resource specification profiling program, RS-Profiler. It is difficult to analytically model effects such as multi-path and component variability. However, models based on measurement data that captures these effects can improve accuracy substantially. We study how to automate the probing process to measure the uncertainty factors and how the resource specification which is constructed on the measured data can be applied to predict the performance of various wireless protocols including routing layer. We validate our method and approach through analytical, and extensive experimental studies. We implement resource specification profiler which collects RF data of all links and all nodes, e.g. RSSI, SNR, PRR, noise floor. We provide two performance prediction algorithms that accurately predict the expected performance of protocols using the cumulative routing metric (e.g. ETX) and the 1-hop routing metric (e.g. $PRR \times D$) based on the collected RF resource specification. These two
algorithms can be applied to any topologies. We prove that performance repeatability within WSN testbed and also across WSN testbed with performance prediction algorithms and RF resource specification with extensive experiments over 18 different 2-dimensional grid networks.

However these data driven methods incur scalability problems in data collection. We present a study on time complexity of RS-Profiling, which is $O(N)$ with the number of nodes $N$. Because the profiling time will be impractically long (e.g. 3 days for full RSSI, noise floor, and PRR for all 16 channels, for all 8 different transmission power levels, for all links of all nodes), we present studies of three methods to relieve this RS-Profiling time scalability problem. First method is a reduction and prediction method on transmission power levels. We initially measure the received signal strength (RSS) of links transmitted with the highest transmission power. For other 7 lower transmission power levels, we predict the received signal strength (RSS) of links based on the RSS values with the measured RSS values transmitted with the highest transmission power. Second method is a reduction method on RSS measurement time. We execute a full profiling without any reduction as a reference or a prediction base for later regular time measurements. In regular time measurements, we predict RSS values of links based on the only noise floor measurement. We compare the difference of noise floor values for each nodes between the initial values and the regular time values, then reflect the difference on the regular time RSS prediction. Because, noise floor measurement can be done in parallel, the profiling time complexity can be reduced to $O(1)$, which is 3 min in our study. Third method reduces RSS measurement time by predicting PRR based on the measured RSS and noise floor. After initial full measurement, we adjust the SNR vs. PRR equation with measurement data. Then, for the regular time profiling, we predict PRR of links with the
measured RSS and noise floor instead of measuring PRR directly. This method can reduce the RSS measurement time to 1% of the full measurement.

**Performance variance.** Performance repeatability of wireless protocols across environments can be achieved not only by matching the expected performance, but also by achieving the comparable performance variances together. Even though the expected performances match across two different networks, if the performance variances doesn’t match, we may not be able to say those two networks have comparable performances. For example, the number of transmissions in a network W1 varies $1 \sim 10$, and the average is 5. The number of transmissions in another network W2 varies $4 \sim 6$, and the average is also 5. Then, even though we can match the expected performance, but in many cases, we may not have the comparable performances due to the unmatched performance variance. Therefore, we study methods to predict the variance of WSN protocols with mathematical equation and with Monte-Carlo simulation. We also study the upper/lower bounds of protocol performance in a given communication environment. We validate our protocol variance study with experimental results.
CHAPTER 2
LINK USAGE SPECTRUM AND CONSISTENT REPRODUCTION OF WIRELESS ROUTING PROTOCOL PERFORMANCE

In this chapter, we present a study answering the questions, “What are the factors impacting performance metric? Given a target network and a test network which are identical modulo of their values for some factors, can the test network be controlled to have, with high probability, the same metric value as the target network?” We consider the problem of obtaining comparable protocol performance when the test and deployment environments differ in RF propagation environment and/or inter-node spacing. To achieve comparable protocol behavior in the two settings, we propose the concept of “link usage spectrum”. Based on the hypothesis that the link usage spectrum is a gross predictor for network performance, we show how to replicate in the test setting the link usage spectrum of the protocol that is expected in the deployment setting. We show our technique for achieving comparable protocol behavior via experiments and simulations in multiple indoor and outdoor propagation environments. The link usage spectrum is protocol specific; we illustrate for a family of protocols how the link usage spectrum is calculated analytically, from the protocol metric for choosing forwarding links in the network, and how power scaling can be used to match the link usage spectrum across networks.
2.1 Link usage spectrum and network transplant error

Wireless network behavior is substantially influenced by the performance of the wireless links between the nodes of the network. Performance of a wireless link between a transmitter and its receivers is determined by the RF channel between the terminals (environment model) and the bit-error-performance of their wireless transceivers (radio model). RF channel models describe the probabilistic relation between link distance and path loss. Specifically, in any given network, links that have the same length experience different channel realizations, due to spatial variations in obstructions and reflectors in the scene. As a result, the received signal strength experienced on links of length $d$ is a random variable $R(d)$. The RF channel model induces a distribution on $R(d)$. For instance, the log-normal shadowing model, a large scale fading model employed commonly in indoor and outdoor link studies, describes the received signal strength as:

$$R(d) = P_t - PL(d_0) - 10\eta \log(d/d_0) + N_\sigma$$

(2.1.1)

where $\eta$ is the path loss exponent, $P_t$ is the transmission power, and $PL(d_0)$ is the path loss observed at distance $d_0$ in dB and $N_\sigma$ is a zero-mean Gaussian random variable with standard deviation $\sigma$, representing spatial variations in the RF environment. The SNR at the receiver $y(d)$ is given by the received signal power $R(d)$ reduced by the noise power $P_0$:

$$y(d) = R(d) - P_0 \text{ (in dB)}$$

(2.1.2)

The radio receiver performance can be characterized by representing the packet reception rate $\text{PRR}(y)$ as a function of the received SNR, $y$. $\text{PRR}(y)$ gives the probability that a packet received with SNR of $y$ will be decoded correctly by the receiver. The relation between the packet-reception-rate and the SNR depends on the modulation scheme and the packet encoding scheme employed by the radio transceiver. The
function $PRR(y)$ is a monotonically increasing function with range $[0, 1]$ and acts as a soft limiter.

The combination of the environment and radio model completely describes the link properties observed in a wireless network for low-rate/time division access where the interference is not significant. Experimental [36] and analytical [37] studies of low power wireless links have shown that in such settings a high percentage of network links will be either good or bad, $> 90\%$ and $< 10\%$, respectively PRR.

![Figure 2.1: PRR distribution of links for indoor and outdoor propagation environments at various transmission power levels (C:Chosen, A:All)](image)

We note that the link reliability statistics reported by previous studies are based on the *a priori* distribution of the link realizations. If we consider the *posterior* distribution of the links that are selected by a given network protocol, the distribution will be skewed towards high PRR values. Figure 2.1 shows the expected PRR of links of various length based on whether or not they were chosen by the forwarding protocol.
We observe that the expected PRR of the links chosen by the protocol is uniformly high and markedly different from the expected PRR of all links at a given distance, especially in the case of long links. As a result, network forwarding performance is grossly determined by the link lengths that are being utilized and less so by the small variations in link qualities. Therefore, in this chapter, we focus on a particular network statistic called the *link usage spectrum*, defined as follows.

**Definition 1.** *Link Usage Spectrum* is the probability distribution with which the network protocol selects links of different length from among all the available links in the network at hand.

*Example 0.* To illustrate the definition of link usage spectrum, consider a wireless network $W = \{(l_j), \eta, \sigma\}$ with link set $\{l_j\}_{j=1}^{M}$ and the RF environment $(\eta, \sigma)$ employing a network protocol $P$. We note $W$ is a probabilistic object, referring to the ensemble of link set realizations. For each realization of the wireless network $W$, network protocol, $P$ chooses a subset of the link set for forwarding of data.

For a one-dimensional linear networks with uniform node spacing, where the link lengths $c_j \equiv c(l_j)$ are constrained to the finite set $\{\tau, 2\tau, 3\tau, \ldots, N\tau\}$, where $\tau$ is the minimum node spacing, the link usage spectrum $L(W, P, i)$ is the discrete probability distribution over the length of links induced by the network protocol $P$, which can be expressed as

$$L(W, P, i) = \text{Prob}[c(l) = i\tau] \quad (2.1.3)$$

*Example 1.* We extend the previous example from a one-dimensional grid to a two-dimensional grid network with uniform node spacing. In this case, the link lengths $c_j \equiv c(l_j)$ and the progress to destination $d_j \equiv d(l_j)$ are constrained to the finite set $\{(a\tau, b\tau) : 1 \leq a, b \leq N\}$, where $\tau$ is the minimum node spacing. We index elements of the above finite set from 1 to M. The link usage spectrum $L(W, P, i)$, then, is the
discrete probability distribution over the length of links and the progress of links to the destination induced by the network protocol $P$.

$$L(W, P, i) = \text{Prob}[c(l) = c_i, d(l) = d_i]$$  \hspace{1cm} (2.1.4)

, where $i$ is the index of $(m\tau, n\tau)$, $c_i = \sqrt{(m\tau)^2 + (n\tau)^2}$, $d_i = c_i \cdot \cos(|\tan^{-1}(\frac{n}{m}) - \theta|)$, $\theta$ is the angle between source and destination.

Note that the link usage spectrum is distinct from the connectivity graph. The links used by the forwarding protocol will be in general only a small subset of the connectivity graph found by maximizing the routing metric over the set of all valid links. The fundamental importance of link usage spectrum stems from the fact that many network-wide metrics can be calculated as averages over link realizations weighted by the usage spectrum, as discussed in Section 1.5. As a result, we propose to use the link usage spectrum to match protocol behavior across scales and environments.

2.1.1 Comparing protocol performance in two settings

Next, we propose a procedure transplanting a network protocol to a different RF environment and/or a different inter-node separation scale. The basic idea is to minimize the distance between the link usage spectra of the two networks; one way of realizing this is by optimizing the selection of the transmission power.

**Definition 2.** Consider a wireless network $W$ with inter-node distances $\{d_j\}_{j=1}^m$ and its scaled version $\tilde{W}$, with inter-node distances $\{\tilde{d}_j = \alpha d_j\}_{j=1}^m$ in RF environments characterized by log-normal scale model parameters $(n, \sigma)$ and $(\tilde{n}, \tilde{\sigma})$ respectively. We define the Transplant Error of a protocol $P$ across the two networks as:

$$\text{XPlantError}(W, \tilde{W}, P) = \sum_{i=1}^n |L(W, P, i) - L(\tilde{W}, P, i)|$$

$\text{XPlantError}$ is essentially the $l_1$ distance between the link usage spectra for the two networks $W$ and $\tilde{W}$ which differ in scale and RF environment. Theorem 1, 2
show the relation between the link usage spectrum and the performance metric. If we can control the XPlantError within a threshold value, we conjecture that the protocol performance will be similar across scale and environment. We assume that transmission power in the network $\tilde{W}$ is variable through $\tilde{P}_t = \tilde{P}_0 + \beta$, where $\beta$ is the power attenuation or amplification in the scaled network. Since the scaled vector in general reduces the node distances for convenient testing, $\beta$ in general is a negative value indicating power attenuation. As a result, the scaled network realizations $\tilde{W}(\beta)$ depends on $\beta^1$. The optimal power attenuation is then chosen to minimize the XPlantError metric:

$$\beta_{\text{opt}} = \arg\min_{\beta} \text{XPlantError}(W, \tilde{W}(\beta), \mathcal{P})$$

(2.1.5)

As shown in [29], if the two networks are in the same environment (i.e., $\eta = \tilde{\eta}$ and $\sigma = \tilde{\sigma}$) then $\beta$ can be chosen such that the link SNR realizations $y_j$ and $\tilde{y}_j$ are samples from the same multivariate Gaussian probability distribution. As a corollary, the optimal $\beta$ in that case would result in identical link usage spectra. For networks in different RF environments, the distribution of link SNR realizations cannot be matched for all link lengths simultaneously and alternative techniques as shown above are required to achieve comparable network behavior. In the next sections, we validate the proposed measure of similarity through experimental and simulation studies using linear and grid networks of wireless nodes employing 802.15.4 radios.

$^1$We could also introduce spatial variations in $\beta$ across the nodes to influence the width of the resulting spectrum, for a better match.
2.2 Experimental study of protocol transplantation

In this section, we present an experimental study of reproducing protocol performance across different WSN environments. We focus our experiments on a well-known tree-based convergecast protocol, called the Collection Tree Protocol (CTP) [6]. In the following, we first give a detailed description of our experimental setup and physical, link and messaging layer assumptions. Then, we study a scaling exercise with a linear network topology in multiple indoor and outdoor settings using data from field experiments. We show that comparable protocol performance is achieved for various metrics (end-to-end delay, mean hop length) if the transmission powers are chosen to minimize the distance between link usage spectra for test and deployment environments.

2.2.1 Experimental setup

We set up a one-dimensional linear topology with a total of 20 TelosB sensor nodes [34]. Each node is separated by 3 ft and elevated about 4 inches from the ground. The TelosB mote platform is equipped with a CC2420 radio [13] and provides eight different transmission power levels: 31(0 dBm), 27(-1 dBm), 23(-3 dBm), 19(-5 dBm), 15 (-7 dBm), 11(-10 dBm), 7(-15 dBm), 3(-25 dBm) [34]. Using a -3dB attenuator in conjunction with software power level settings, 15 distinct transmission power levels can be realized.

We performed experiments in three settings, to capture different environment properties. We used a single output level of -13dBm for an indoor network in a long office corridor; three output levels (-6 dBm, -3 dBm and 0 dBm) for an outdoor network in a parking lot, corresponding to power scaling coefficient $\beta$ of (7 dB, 10 dB, 13 dB); and three output levels (-18 dBm, -13 dBm and -10 dBm) for an indoor network in an open warehouse.
2.2.2 Physical & link layer

The CC2420 [13] radio is compatible with the 2.4GHz 802.15.4 standard. The 802.15.4 physical layer employs block direct-sequence spread spectrum code with 2MChip/s chip rate and 250 kbps data rate to achieve processing and coding gain. The transmitter modulates the carrier using offset quadrature phase shift keying (O-QPSK) with half-sine shaping which is equivalent to minimum shift keying (MSK) modulation, which has the following bit error rate (BER):

\[
\text{BER} = Q(\sqrt{2y/\text{PG}/\text{CG}}) = \frac{1}{2}\text{erfc}(\sqrt{y/\text{PG}/\text{CG}}) \quad (2.2.1)
\]

where \(y\) gives the SNR. The processing gain (PG) for 802.15.4 is given by \(10 \log(2/0.25) = 9\) dB. The coding gain (CG) depends on the increased Hamming distance between the codes and is a function of the SNR itself. For a low packet error rate region, the coding gain is approximated as 2 dB [22]. Thus, the Packet Reception Rate equation is calculated to be

\[
\text{PRR} = \left(1 - \frac{1}{2}\text{erfc}(\sqrt{(x - 11 - P_0 \text{ dB})})\right)^{8\times\text{packet\_size}} \quad (2.2.2)
\]

where \(x\) denotes the SNR in dBm. We note that the bit-error-rate approximation given in Equation 3.2.1 assumes coherent demodulation using carrier phase information. Practical transceiver designs use non-zero IF and noncoherent demodulation. The non-ideal receiver structures can be approximated with SNR reduction or equivalently increase in the noise floor \(P_0\). We use a radio sensitivity of -94 dBm to adjust for the noise power \(P_0\).

To complete our analytical model, we require RF environment parameters: Path Loss Exponents (PLE) of indoor and outdoor and standard deviation of RSSI(dB). We performed RSSI measurements in a corridor in the second floor of the Dreese Lab Building and outdoor in a parking lot. For the indoor tests, we measured RSSI at 20 different distances (1 ~ 20 unit (1 unit = 3ft)) within maximum communication range.
range with the highest transmission power level (0dBm). For the outdoor test, 10 measurements were taken from the distance of 1 ∼ 10 unit distances where 30 ft seems to be the maximum communication range with the same transmission power, 0 dBm. Figure 2.2 shows the observed received signal values and the associated log-normal fit. Table 3.1 presents the summary.

Figure 2.2: Indoor corridor, outdoor, and indoor warehouse (RSSI vs. distance)

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Indoor Corridor</th>
<th>Outdoor</th>
<th>Indoor Warehouse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Path Loss Exponent</td>
<td>1.7555</td>
<td>2.2776</td>
<td>1.3545</td>
</tr>
<tr>
<td>RSSI Standard Dev.</td>
<td>5.0 dB</td>
<td>4.2 dB</td>
<td>5.0 dB</td>
</tr>
</tbody>
</table>

Table 2.1: Log normal model variables for indoor corridor, outdoor, and indoor warehouse RF environments
2.2.3 Messaging layer

CTP [9] is a tree-based collection protocol. Nodes generate routes to the sink using a routing gradient. CTP uses ETX [15] as the default routing metric. ETX implicitly favor long links over short links because each node selects the path with the minimum number of expected transmissions. Therefore, we expect ETX works similar to other metrics which give preference to long links, such as PRR×d [31].

For the linear network with 20 nodes, Node 1 is designated as the source, and Node 20 is set as the sink. All other intermediate nodes act as multihop relays. A source packet is generated every two seconds. The low data rate ensures no interference from previous packets sent through the network. For each power level we use 1000 source packets and log all the paths that each packet have gone through, and only exclude the first and lost hop to avoid edge effects. Using the information embedded into the packets we compute (i) median link length, (ii) end-to-end latency, and (iii) link usage spectrum, for each power level.

2.2.4 Results

We first present link usage spectrum for indoor and outdoor environments at various transmission power levels using data from the experiments in Figure 2.3. We see that the closest link usage spectra is observed for the corridor and the outdoor experiments when the outdoor network employs a transmission power level between -3dBm and 0dBm, and for the corridor and warehouse experiments when the warehouse network uses a transmission power level between -18dBm and -13dBm. Table 2.2 shows that the outdoor XPlantError is minimized for the 0dBm case, and the warehouse XPlantError is minimized for the -18dBm case.

Next, we show that matching link usage spectra results in consistent protocol performance as measured by commonly adopted metrics of mean-link-length and
Figure 2.3: Cumulative link usage spectrum. (1) indoor corridor vs outdoor, (2) indoor corridor vs warehouse

Table 2.2: Comparison of XPlantError, $l_1$ distance, between corridor network with tx=-13dBm and various outdoor (O)/warehouse (W) networks

<table>
<thead>
<tr>
<th>$l_1$ distance</th>
<th>O -6dBm</th>
<th>O -3dBm</th>
<th>O 0dBm</th>
<th>W -10dBm</th>
<th>W -13dBm</th>
<th>W -18dBm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anlytical</td>
<td>2.75</td>
<td>1.87</td>
<td>1.20</td>
<td>3.24</td>
<td>1.69</td>
<td>1.59</td>
</tr>
</tbody>
</table>

end-to-end latency in these experiments. Table 2.3 and 2.4 shows summary statistics for link length and end-to-end delay for the indoor and outdoor environments, for each specified transmission power level. We see that the performance of the corridor environment network is best matched by the outdoor environment network between when the latter uses a transmission power level between -3 dBm and 0 dBm, and by the warehouse environment network when the latter uses a transmission power level between -18 dBm and -13 dBm, which is consistent with our main hypothesis.
In this section, towards showing one way in which the link usage spectrum in a test environment can be calculated, we derive analytical expressions that are accurate approximations of the link usage spectrum. Our derivation is specific to forwarding protocol that maximizes \( \text{PRR} \times d \) protocol metric under lognormal shadowing model; this protocol metric is sometimes referred to as ELD. With such an analytical model, we can choose the attenuation level for matching link usage spectrum of two different deployment settings, so that comparable network performance may be achieved.

We also present simulation (specifically, Monte Carlo and TOSSIM) results that corroborate the analytical framework developed here.

We note that while the results here are presented using \( \text{PRR} \times d \) as the protocol
metric for evaluating links, they are easily customized to other forwarding protocols based on optimization of network metrics encapsulating PRR, SNR and \( d \).

### 2.3.1 Analytical model for one-dimensional and two-dimensional uniform graphs

The link usage statistics represent order statistics for the chosen protocol metric. It is straightforward to express the probability of attaining the maximum value among collection of random variables using the probability and cumulative density functions of the underlying random variables. However, for many transport protocols, the protocol metric for each link has a different probability distribution with non-identical support. This imposes a complex partition of the underlying multi-dimensional of link SNR’s with nonlinear boundaries prohibiting formulation of closed form expressions for link usage statistics. Instead of relying on computationally expensive numerical methods for calculating the relative volume of each partition, we derive analytical expressions for link usage statistics by approximating the partition boundaries with piecewise linear functions resulting in simple, accurate expressions for link usage. These provide computational savings over direct numerical integration method and, more importantly, allow generalization for the asymptotic case of the network size going to infinity.

**Theorem 1.** For a protocol \( \mathcal{P} \) that uses \( PRR \times d \) as the metric for choosing forwarding links, the probability of choosing link \( l_i \) over link \( l_j \) is expressed as follows:

\[
P( PRR(y_i) \cdot d_i > PRR(y_j) \cdot d_j )
\]

\[
\simeq \int_{-\infty}^{a} \frac{1}{\sigma \sqrt{2\pi}} \left( \frac{1}{2} + \frac{1}{2} \text{erf} \left( \frac{y_i - \mu_{y_j}}{\sqrt{2}\sigma} \right) \right) e^{-\frac{(y_i - \mu_{y_j})^2}{2\sigma^2}} dy_i 
\]

\[
+ \int_{a}^{\infty} \frac{1}{\sigma \sqrt{2\pi}} \left( \frac{1}{2} + \frac{1}{2} \text{erf} \left( \frac{g(\beta_j) - \mu_{y_i}}{\sqrt{2}\sigma} \right) \right) e^{-\frac{(y_i - \mu_{y_i})^2}{2\sigma^2}} dy_i,
\]
where \( y_i, y_j \) are SNR (dB) values of link \( l_i, l_j \), and \( d_i, d_j \) are progresses to destination for link \( l_i, l_j \), \( \beta_j = \frac{d_j}{d_i} \), \( a = \text{PRR}^{-1}(\min\{\beta_j, \frac{1}{\beta_j}\}) \), \( g(\beta_j) = a \) if \( \beta_j \geq 1 \) and \( g(\beta_j) = \infty \) if \( \beta_j < 1 \).

**Proof.** The probability of choosing \( l_i \) over \( l_j \) with metric \( \text{PRR} \times D \),

\[
\text{Prob}(\text{PRR}(y_i) \cdot d_i > \text{PRR}(y_j) \cdot d_j) = \text{Prob}(\text{PRR}(y_i) > \beta_j \cdot \text{PRR}(y_j))
\]

where we define the link length ratio as \( \frac{d_j}{d_i} = \beta_j \). Without loss of generality, we can assume \( \beta_j < 1 \). Then we can approximate the boundary between the protocol metrics calculated for two links as:

\[
\text{PRR}(y_i) > \beta_j \cdot \text{PRR}(y_j) \approx y_i > h(y_j)
\]

, where

\[
h(y_j) = \begin{cases} y_j, & \text{if } y_j < a \\ a, & \text{if } y_j \geq a \end{cases}
\]

, when \( a = \text{PRR}^{-1}(\beta_j) \)

\[
\text{Prob}(\text{PRR}(y_i) > \beta_j \cdot \text{PRR}(y_j)) \approx \text{Prob}(y_i > y_j, y_j < a) + \text{Prob}(y_i \geq a, y_j \geq a)
\]

\[
= \text{Prob}(y_i > y_j, y_i < a) + \text{Prob}(y_i \geq a)
\]

Because \( y_i, y_j \) are Gaussian:

\[
\text{Prob}(\text{PRR}(y_i) > \beta_j \cdot \text{PRR}(y_j)) \approx \int_{-\infty}^{a} \frac{1}{\sigma \sqrt{2\pi}} \left( \frac{1}{2} + \frac{1}{2} \text{erf} \left( \frac{y_i - \mu_y}{\sigma \sqrt{2}} \right) \right) e^{-\frac{(y_i - \mu_y)^2}{2\sigma^2}} dy_i
\]

\[
+ \int_{a}^{\infty} \frac{1}{\sigma \sqrt{2\pi}} \left( \frac{1}{2} + \frac{1}{2} \text{erf} \left( \frac{g(\beta_j) - \mu_y}{\sqrt{2}\sigma} \right) \right) e^{-\frac{(y_i - \mu_y)^2}{2\sigma^2}} dy_i,
\]

25
where

\[
g(\beta_j) = \begin{cases} 
  a, & \text{if } \beta_j \geq 1 \\
  \infty, & \text{if } \beta_j < 1 
\end{cases}
\]

\[L(W, P, \text{index}(l)) = P\left[\text{PRR}(y(l)) \cdot d(l) = \max_{i=1..M} \{\text{PRR}(y_i) \cdot d_i\}\right]
\]

\[
\simeq \sum_{k=0}^{M-1} \int_{a_k}^{a_{k+1}} \frac{1}{\sigma \sqrt{2\pi}} \left[\prod_{i=1}^{k} \left(\frac{1}{2} + \frac{1}{2} \text{erf}\left( \frac{g(\beta_i) - \mu y_i}{\sqrt{2\sigma}} \right) \right)\right] \left[\prod_{i=k+1}^{M} \left(\frac{1}{2} + \frac{1}{2} \text{erf}\left( \frac{y(l) - \mu y_i}{\sqrt{2\sigma}} \right) \right)\right] e^{-\frac{(y(l) - \mu y(l))^2}{2\sigma^2}} dy(l),
\]

where \(\beta_i = d_i/d(l)\), \(a_i = \text{PRR}^{-1}(\min\{\beta_i, \frac{1}{\beta_i}\})\), and the links are enumerated such that \(a_1 \leq a_2 \leq ... \leq a_{M-1}\), with \(a_0 = -\infty, a_M = \infty\) and \(y_i\) is the SNR experienced by link \(l_i\), and

\[
g(\beta_i) = \begin{cases} 
  a_i, & \text{if } \beta_i \geq 1 \\
  \infty, & \text{if } \beta_i < 1 
\end{cases}
\]

\[L(W, P, \text{index}(l)) = P\left(\text{PRR}(y(l)) \cdot d(l) = \max\{\text{PRR}(y_i) \cdot d_i : 1 \leq i \leq M\}\right)
\]

\[= P(\text{PRR}(y(l)) \cdot d(l) > \text{PRR}(y_1) \cdot d_1) \land \ldots \land (\text{PRR}(y(l)) \cdot d(l) > \text{PRR}(y_M) \cdot d_M))
\]

Theorem 2. For a protocol \(P\) which uses \(\text{PRR} \times d\) as the metric for choosing forwarding links, the probability of choosing link \(l\) of length over all other links is expressed as follows:
Let’s put
\[
f(\cdot) = \frac{1}{(\sigma \sqrt{2\pi})^M} e^{-\frac{(y_1 - \mu_{y_1})^2 + \ldots + (y_M - \mu_{y_M})^2}{2\sigma^2}}
\]

Because the conjunctions of the conditions will be translated into multiple integrals.

Apply Theorem 1 here, we have
\[
L(W, P, index(l)) \simeq \int_{-\infty}^{a_1} \int_{-\infty}^{y(l)} \ldots \int_{-\infty}^{y(l)} f(\cdot) dy_M \ldots dy_1 dy(l)
\]
\[
+ \int_{a_1}^{a_2} \int_{-\infty}^{g(\beta_1)} \int_{-\infty}^{y(l)} \ldots \int_{-\infty}^{y(l)} f(\cdot) dy_M \ldots dy_1 dy(l)
\]
\[
+ \int_{a_2}^{a_3} \int_{-\infty}^{g(\beta_1)} \int_{-\infty}^{g(\beta_2)} \int_{-\infty}^{y(l)} \ldots \int_{-\infty}^{y(l)} f(\cdot) dy_M \ldots dy_1 dy(l)
\]
\[
+ \ldots
\]
\[
+ \int_{a_n}^{\infty} \int_{-\infty}^{g(\beta_1)} \ldots \int_{-\infty}^{g(\beta_n)} f(\cdot) dy_M \ldots dy_1 dy(l)
\]

Then, we have
\[
L(W, P, index(l)) \simeq \int_{-\infty}^{a_1} \frac{1}{\sigma \sqrt{2\pi}} Z_1 \ldots Z_M e^{-\frac{(y(l) - \mu_{y(l)})^2}{2\sigma^2}} dy(l)
\]
\[
+ \int_{a_1}^{a_2} \frac{1}{\sigma \sqrt{2\pi}} \Phi_1 Z_2 \ldots Z_M e^{-\frac{(y(l) - \mu_{y(l)})^2}{2\sigma^2}} dy(l)
\]
\[
+ \ldots
\]
\[
+ \int_{a_{n-1}}^{a_n} \frac{1}{\sigma \sqrt{2\pi}} \Phi_1 \ldots \Phi_{M-1} Z_M e^{-\frac{(y(l) - \mu_{y(l)})^2}{2\sigma^2}} dy(l)
\]
\[
+ \int_{a_n}^{\infty} \frac{1}{\sigma \sqrt{2\pi}} \Phi_1 \ldots \Phi_M e^{-\frac{(y(l) - \mu_{y(l)})^2}{2\sigma^2}} dy(l)
\]

, where \( Z_i = \left( \frac{1}{2} + \frac{1}{2} \text{erf} \left( \frac{y(l) - \mu_{y_i}}{\sqrt{2\sigma}} \right) \right) \), \( \Phi_i = \left( \frac{1}{2} + \frac{1}{2} \text{erf} \left( \frac{g(\beta_i) - \mu_{y_i}}{\sqrt{2\sigma}} \right) \right) \), \( y(l), y_i \)

are SNR (dB) values of link \( l, l_i, index(l) \neq i \)

In summary,
\[
L(W, P, index(l))
\]
\begin{align*}
&\sum_{k=0}^{M} \int_{a_k}^{a_{k+1}} \frac{1}{\sigma \sqrt{2\pi}} \left[ \prod_{j=1}^{k} \left( \frac{1}{2} + \frac{1}{2} \text{erf} \left( \frac{g(\beta_j) - \mu y_j}{\sqrt{2} \sigma} \right) \right) \right] \\
&\left[ \prod_{j=k+1}^{M} \left( \frac{1}{2} + \frac{1}{2} \text{erf} \left( \frac{y_j(l) - \mu y_j}{\sqrt{2} \sigma} \right) \right) \right] e^{-\frac{(y(l) - \mu y(l))^2}{2 \sigma^2}} \, dy(l)
\end{align*}

The proof is an immediate extension of Theorem 1. Both theorems rely on the approximation used in Equation 2.3.1 to convert the nonlinear boundary where one link is preferred over another to a piece-wise linear boundary suitable for close-form integration.

2.3.2 Validation of analytical approximations

We performed Monte-Carlo simulations to verify the accuracy of the analytical approximation of the link usage spectrum for various radio propagation environments and transmission power levels. Figure 3.1 shows the analytical and Monte Carlo simulation results of link usage spectra for indoor and outdoor environments with specified transmission powers. The Monte-Carlo simulation results are nearly identical to the analytical results, supporting the approximations used in the derivation of the analytical expressions of link usage spectrum.

We also compared the results of the analytical model with the observed link usage spectrum using the experiments described in the previous section. We observe that although the general behavior of link usage spectrum as a function of transmission power is consistent between analytical and experimental results, the links chosen in the field experiments are in general shorter than the analytically derived link lengths. Careful analysis of the temporal variations in the experimentally observed tree structure suggests that temporal RSSI fluctuations are one potential reason for the observed gap. The analytical model derived in this section is a large scale fading
model and does not consider the impact of temporal RSSI variations on link usage pattern. As shown in Figure 2.1, if we only consider spatial RSSI variations, long term average PRR of all the chosen links are close to unity. However, if we consider temporal RSSI variations, long links will suffer in greater degree because they are likely to be closer to the sensitivity threshold of the receiver, which in turn will result in major fluctuations in their PRR values. Consequently, the routing protocol will deselect over time long links encountering temporal fades. To account for this observed behavior, we reject links whose unconditional (prior) expected PRR is larger than 1%. As an example, for Figure 2.1, for an outdoor network with 0 dB transmission power, we will reject links whose length is less than or equal to 10 hops. The result is shown in Figure 2.5.

**Two-dimensional simulation study.** In this section we present a comparison of analytical and simulation results of the link usage spectrum for the CTP protocol for a two-dimensional 10×10 grid topology. The internode distance is set to 6 ft, the source node is located at the leftmost bottom corner, and the sink node is at the rightmost top corner. First, using the analytical expression given in Theorem 2 we
Figure 2.5: Cumulative link usage spectrum ((1) Theorem 2 with refinement, (2) experiments)

calculate the two-dimensional link usage spectrum, and then compare it to empirical link usage data obtained through the TOSSIM 2 [7] simulator.

Figure 2.6: Link usage spectrum (analytical). (1) indoor -17dBm and (2) outdoor -5dBm

Figure 2.6 shows the analytical results calculated. Figure 2.7 shows the TOSSIM 2 [7] simulation results. We note that analytical expressions assume an infinite network model, calculate performance for a single source node, and scale the results to
a finite size network. Analytical results show that the indoor network with transmission power (-17dBm) is matched optimally with the outdoor network with (-5dBm) transmission power, using the XPlantError metric. Figure 2.6 shows contour plots for the analytically derived two-dimensional link usage spectra for indoor (-17dBm) and outdoor (-5dBm). We observe analytically that the link usage is very similar at the optimal matching attenuation of $-12\, dB$.

The simulation results of these two networks closely follow the analytical results. In Figure 2.7, the cumulative usage weights of indoor -17dBm matches well with outdoor -5dBm, and usage weights peak around progress 3 and saturate around progress 5, which conforms with Figure 2.6, where usage weights peak at coordinate (2,2) (progress $\simeq 3$) and saturate at coordinate (3.5,3.5) (progress $\simeq 5$). We also give
cumulative usage weights for a second outdoor network of transmission level of 0 dB, as an example of mismatch of usage weights.

Table 2.5 shows average end-to-end delays and link progresses for each environment. As predicted by the analytical results, indoor -17dBm matches outdoor -5dBm better than it does 0dBm.

2.4 Two-dimensional experimental study across testbeds

In this section, we present the results of an experimental study with two-dimensional 4 × 4 grid topology across two testbeds. We experiment with CTP in two well-known indoor WSN testbeds: KanseiGenie [2] and Motelab [3].

2.4.1 Experimental setup

First we setup a two-dimensional grid of 16 TelosB sensor nodes in the KanseiGenie testbed on a 4 × 4 grid topology. The selected KanseiGenie nodes are separated by 10 ft and elevated about 4 ft from the ground. Each node consists of a TelosB mote platform equipped with a CC2420 radio and a -3dB attenuator. In the Motelab, the two-dimensional 4 × 4 grid topology with equal spacing can only be achieved approximately, since the Motelab nodes are randomly spread across the testbed. We carefully chose 16 nodes to form an approximate regular grid with an average spacing of 14 ft. The Motelab nodes are TMoteSky mote platforms equipped with CC2420 radio. There are three major differences between the networks implemented at the KanseiGenie and the Motelab. First, TelosB nodes of the KanseiGenie are equipped with dipole antennas, whereas TMoteSky nodes of the Motelab have printed circuit board (PCB) antennas. Therefore, the KanseiGenie network is expected to have better RF propagation performance at the same power level. Second, node separation of the KanseiGenie 4 × 4 grid is about 10 ft, but that of the Motelab is approximately
14 ft. Third, the KanseiGenie has no walls or occlusions, but the Motelab network is installed in a building with many rooms and corridors. In summary, we expect the radio propagation environment of the two testbeds to differ substantially. To evaluate these differences, the RF propagation characteristics of the two testbeds are measured, and shown in Figure 2.8. Assuming a log-normal fit, we obtained RF model parameters and used the proposed transplanting protocol method of matching $l_1$ distance of Link Usage Spectra to repeat the protocol performance of the Motelab network in the target KanseiGenie network.

![Figure 2.8: KanseiGenie and Motelab in channel 26 (RSSI vs. distance)](image)

We used a single output level of -15dBm for the network in Motelab and four output levels (-15 dBm, -10 dBm, -5 dBm and 0 dBm) for the network in KanseiGenie, corresponding to power scaling coefficient $\beta$ of (0dB, 5dB, 10 dB, 15 dB).
2.4.2 Message layer

In this two-dimensional testbed experiment study, we employ a modified version of CTP that uses PRR×d as the routing metric. The CTP with PRR×d chooses the best link which has the highest PRR×d metric value among the candidate links. In the experiments, all nodes have the geographic information (coordinates of all nodes) to calculate \(d\) (progress to destination) among the available links.

The source node is located at grid location (1,1) and sends packets over the multihop network to the sink node located at (4,4). All other intermediate nodes act as multihop relays. A source packet is generated every two seconds. We have chosen this low data rate to assure no self interference from previous packets sent through the network. For each power level we use 1,000 source packets and log the entire path that each packet goes through, and then exclude the first and lost hop to avoid edge effects. Using the information embedded into the packets we compute three performance metrics at each power level: (i) link progress to destination, (ii) end-to-end latency, and (iii) link usage spectrum.

2.4.3 Results

Based on log-normal fit to the RF propagation environments shown in Figure 2.8 and Theorem 2, we analytically calculate the Link Usage Spectra of the CTP with transmission power of -15dBm at Motelab, and those of the CTP with transmission power of -15dBm, -10dBm, -5dBm, and 0dBm. Figure 2.9 (1) depicts the analytical results. The link usage spectrum of the CTP with -15dBm at the Motelab has the closest \(l_1\) distance to that of the KanseiGenie with -10dBm. The values of the corresponding \(l_1\) distances are presented in Table 2.6. Figure 2.9 (2) shows the corresponding experimental results. Consistent with our analytical results, the experimental results show
Figure 2.9: Cumulative link usage spectrum. (1) analytical, (2) experimental, K:KanseiGenie, M: Motelab

that the link usage spectrum of Motelab at -15dBm matches closely the observed link usage spectrum of KanseiGenie at -10dBm.

<table>
<thead>
<tr>
<th>$l_1$ distance</th>
<th>KanseiGenie</th>
<th>KanseiGenie</th>
<th>KanseiGenie</th>
<th>KanseiGenie</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytical</td>
<td>1.4529</td>
<td>0.1955</td>
<td>0.8706</td>
<td>1.4931</td>
</tr>
</tbody>
</table>

Table 2.6: Comparison of analytical XPlantError, $l_1$ distance, between the Motelab testbed with tx=-15dBm and the KanseiGenie testbed with various transmission powers

Table 2.7 shows the experimental results for two performance metrics: end-to-end delay and progress to destination. Again as suggested by our analytical results from Table 2.6 and Figure 2.9, experimental results in Table 2.7 shows that the closest performance is achieved at the transmission power pair (Motelab -15dBm and KanseiGenie -10dBm) for both performance metrics (end-to-end delay and link progress).
Table 2.7: End-to-end delay and link progress for two-dimensional environments (testbeds experiments, K:KanseiGenie, M: Motelab)

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>K</th>
<th>K</th>
<th>K</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td>e-2-e delay</td>
<td>1.99</td>
<td>1.52</td>
<td>2.00</td>
<td>4.47</td>
<td>5.55</td>
</tr>
<tr>
<td>Link Progress</td>
<td>3.17</td>
<td>6.04</td>
<td>3.10</td>
<td>1.29</td>
<td>1.10</td>
</tr>
</tbody>
</table>

2.4.4 Discussion on theory vs. experiment gap

While our performance matching method chooses successfully a power attenuation factor $\beta$ in this experimental study, we still observe a gap between the Link Usage Spectra calculated analytically and measured experimentally at each power level. Specifically, as shown in Figure 2.9, experimental Link Usage Spectrum consists of links with shorter link progress than the ones that are analytically predicted for low transmission power levels (-10dBm, -15dBm). We hypothesize two potential explanations for this discrepancy: First, to obtain the analytical results, we assumed that the topology is infinite so that the statistics of the collected chosen links will exhibit the predicted link distribution obtained through aggregating spatially across the network. However, the experiment grid topology has finite extent, and as a result the chosen links are biased towards shorter link progress for lower transmission power cases than predicted by the analytical model which is based on an infinite topology.

Second, for deriving the analytical results we did not consider the effect of RF interference from external networks present in the same frequency band. This RF interference may contribute to the observed gap between theoretical and experimental data. To illustrate the strong role that RF interference can play in received power levels, we collected noise floor measurements with contributions from two interference
Figure 2.10: Noise floor of the KanseiGenie channel 26

Figure 2.11: Noise floor of the KanseiGenie channel 12

sources: cross technology interference (CTI) and interference from other experiments running simultaneously at other nearby nodes in the testbed. The KanseiGenie is located at a warehouse, free from Bluetooth traffic. Therefore nearby WiFi networks will be the only source of CTI. KanseiGenie experiments are conducted typically at Channel 26 of CC2420 radio since none of the WiFi channels overlap with the CC2420 channel 26. However, noise floor measurements given in Figure 2.10 and Figure 2.11 shows that the observed noise floor of channel 26 is quite higher due to simultaneously running experiments, despite the WiFi CTI impact on Channel 12.
The variations in link SNR due to the elevated noise floor are only partly captured by the RF propagation model that aggregates results across nodes and time. In reality, the external interference sources are not distributed uniformly across space and the external traffic varies with time. These variations in link SNR impacts long links since they are likely to operate close to the radio sensitivity threshold, providing a likely reason for the shorter links than expected links observed empirically at lower transmission power levels.

2.5 Conclusion

In this chapter, we proposed a method for achieving comparable wireless protocol performance across deployment and test environments. We have found the matching method to be valid across several protocols (which in turn were based on different forwarding metrics), for various performance metrics, and diverse RF propagation environments.

The work presented in this chapter can be extended in several directions. First, the proposed analytic method for predicting link usage provides a good approximation for networks with links of limited temporal variation, but refinements are required to handle networks with links of significant temporal variation. Second, error metrics other than XPlantError can be employed for matching performance. One idea is to adjust the transmission power attenuation to match the expected value of a single specified performance metric, such as at SNR, PRR, End-to-end Latency \((1/PRR \times D)\), ELD, \(PRR \times D\), as follows:

\[
\left| \sum_{i=1}^{n} g(y_i | l_i \uparrow) L(W, \mathcal{P}, i) - \sum_{i=1}^{n} g(\tilde{y}_i | l_i \uparrow) L(\tilde{W}(\beta), \mathcal{P}, i) \right| \tag{2.5.1}
\]

where \(l_i \uparrow\) means \(l_i\) is the conditional expectation given the link \(l_i\) was chosen. \(g\) is the performance metric chosen by the designer. It is easy to show that this error metric
is bounded above by the $l_1$ distance between the link usage spectra proposed in this chapter. This error metric will be specific for the chosen performance metric, whereas

Figure 2.12: Functional relation between indoor corridor and outdoor transmission power for minimizing $l_1$ distance between the link usage spectra (Red), and for matching end-to-end latency performance (Cyan).

the method we propose based on minimizing the distance between link usage spectra will be a generic method for obtaining comparable performance over a wide variety of network metrics. Figure 2.12 compares the optimal attenuation level for link usage spectrum matching with the optimal attenuation required for matching end-to-end latency of test/deployment networks. We observe that the two approaches yield essentially the same attenuation factors.
CHAPTER 3

ANALYTICAL MODELING OF EXPECTED PERFORMANCE AND PERFORMANCE REPEATABILITY OF LOW POWER WIRELESS SENSOR NETWORK PROTOCOLS

In this chapter, we study the question, “What is the mean and variance of a WSN performance metric? Given a target network and a test network which are identical modulo their value for some factor, can the test network be controlled to have, with high probability, the same performance metric value as the target network?”.

In chapter 2, we defined and mathematically formulate link usage spectrum of wireless protocols for target networks. In this chapter we propose a method to achieve performance repeatability across test and target environments, that relies on analytical prediction of expected protocol performance as a function of RF environment parameters and forwarding protocol. For the validation of the proposed method, we present analytical, simulation, and experimental results for one-dimensional networks deployed in indoor and outdoor propagation environments, and for two-dimensional networks on four major indoor WSN testbeds to validate the performance of the proposed method on achieving repeatable protocol behavior across diverse set of RF environments.
3.1 Wireless protocol performance modeling

In this section we present an analytical tool for determining power scaling to repeat WSN protocol performance across testbed and deployment environments. Wireless protocol performance is determined by the subset of links that are utilized by the wireless protocol and the quality of these chosen links. The quality of the each link in turn is characterized by the RF channel between the terminals (environment model) and the bit-error-performance of their wireless transceivers (radio model).

In the following we derive link usage and link quality statistics for a given network protocol by combining the RF environment and radio receiver models. Then we combine these statistics to assess the expected performance of the wireless protocol.

We consider a wireless network $\mathcal{W} = \{\{l_j\}, \eta, \sigma\}$ with link set $\{l_j\}_{j=1}^M$ and the RF environment $(\eta, \sigma)$, where $\eta$ is the path loss exponent (PLE) and $\sigma$ is the standard deviation of the received signal strength, employing a network protocol $\mathcal{P}$. We note $\mathcal{W}$ is a probabilistic object, referring to the ensemble of link set realizations. For each realization of the wireless network $\mathcal{W}$, network protocol $\mathcal{P}$ chooses a subset of the link set for forwarding of data. We assume $\mathcal{P}$ is a possibly randomized rule for selecting links based on a metric which combines link length $c(l_j)$ and link quality $Q(y(l_j))$ to score the link $l_j$, and the SNR value of the link, $y(l_j)$. As an example we consider two-dimensional grid networks, where the link lengths $c_j \equiv c(l_j)$ and the progress to destination $d_j \equiv d(l_j)$ are constrained to the finite set $\{(a\tau, b\tau) : 1 \leq a, b \leq N\}$, where $\tau$ is the minimum node spacing. We index elements of the above finite set from 1 to $M (= N^2)$. The link usage spectrum $L(\mathcal{W}, \mathcal{P}, i)$ is the discrete probability distribution over the length of links and the progress of links to the destination induced by the network protocol $\mathcal{P}$.

$$L(\mathcal{W}, \mathcal{P}, i) = \text{Prob}[c(l) = c_i, d(l) = d_i] \quad (3.1.1)$$
, where $i$ is the index of $(a, b)$, $c_i = \sqrt{(a\tau)^2 + (b\tau)^2}$, $d_i = c_i \cdot \cos(|\tan^{-1}(\frac{b}{a}) - \theta|)$, $\theta$ is the angle between source and destination. The link usage spectrum is a universal summary statistics of the network behavior. The fundamental importance of link usage spectrum stems from the fact that many network wide metrics can be calculated as the average over link realizations weighted by the link usage spectrum. The link usage spectrum can be calculated in situ empirically as averages over node link usages or can be derived directly from RF environment and physical layer model.

3.1.1 Calculating the expected performance metrics

In this subsection, we derive an equation to calculate the expected performance metric (Theorem 3). Let $f(y_i, d_i)$ be an arbitrary performance metric over the link quality, SNR ($y_i$), and the progress to destination ($d_i$). We assume the performance of a wireless link is measured by $f(y_i, d_i)$. Then the expected performance metric $E[f|W, P]$ is computed as the average over the linkset $l_i$, weighted by the link usage spectrum $L(W, P, i)$ (Theorem 2):

$$E[f|W, P] = \sum_{i=1}^{M} E(f(y_i, d_i|l_i \uparrow))L(W, P, index(l_i))$$

(3.1.2)

Conditional moments of the SNR of the chosen links (Result 4) can be used to approximate $E(f(y_i, d_i|l_i \uparrow))$.

**Theorem 3.** The expected performance metric, $E(f(y_i, d_i|l_i \uparrow))$, of the link $l_i$ can be calculated as below:

$$E(f(y_i, d_i|l_i \uparrow)) \approx f(E(y_i|l_i \uparrow), d_i)$$

$$+ \frac{f''(E(y_i|l_i \uparrow), d_i)}{2} (E(y_i^2|l_i \uparrow) - (E(y_i|l_i \uparrow))^2)$$

, where $y_i$ is SNR (dB) values of link $l_i$, $d_i$ is the progress to destination of $l_i$
Proof. We review the straightforward argument based on the Taylor expansion of the function $f(y, d)$ with respect to its first argument.

$$E[f(y_i|l_i \uparrow)] = E[f(\mu_{y_i} + (y_i - \mu_{y_i})|l_i \uparrow)]$$

Using the Taylor expansion,

$$f(x) = \sum_{k=0}^{n} \frac{f^{(n)}(a)}{n!} (x - a)^n$$

we have

$$\approx E[f(\mu_{y_i}|l_i \uparrow) + f'(\mu_{y_i}|l_i \uparrow)(y_i - \mu_{y_i}|l_i \uparrow)$$

$$+ \frac{1}{2} f''(\mu_{y_i}|l_i \uparrow)(y_i - \mu_{y_i}|l_i \uparrow)^2]$$

$$\approx f(E(y_i|l_i \uparrow)) + \frac{f''(E(y_i|l_i \uparrow))}{2} VAR(y_i|l_i \uparrow)$$

Finally,

$$= f(E(y_i|l_i \uparrow)) +$$

$$\frac{f''(E(y_i|l_i \uparrow))}{2} (E(y_i^2|l_i \uparrow) - (E(y_i|l_i \uparrow))^2$$

Using Theorem 3, we can calculate the most common performance metrics: the expected total number of transmissions and the expected progress to destination. As an example, by choosing $f(y, d) = \frac{1}{PRR(y) \times d}$, we can calculate the expected number of transmissions per unit distance. By multiplying the progress to the destination with the calculated expected delay per unit distance, we can calculate the expected end-to-end number of transmissions of the chosen links.
3.1.2 Calculating the moments of the conditional SNR

In this subsection, we derive an equation (Theorem 4) to calculate the moments of the SNR of a link conditioned on being chosen over all other links. We revisit the link usage spectrum derivation (Theorem 2), which was presented in [20], for further extension in Theorem 4.

**Theorem 4.** When an application uses $\text{PRR} \times d$ as the metric for choosing forwarding links, the expected value of the $m$-th moment of the SNR of a link $l$ conditioned on the fact that it is chosen by the forwarding metric is given by:

$$E(y(l)^m | l \uparrow) \simeq \frac{1}{w} \sum_{k=0}^{M-1} \int_{a_k}^{a_{k+1}} \frac{y(l)^m}{\sigma \sqrt{2\pi}} \left[ \prod_{i=1}^{k} \left( \frac{1}{2} + \frac{1}{2} \text{erf} \left( \frac{g(\beta_i) - \mu_i}{\sqrt{2}\sigma} \right) \right) \right]$$

$$\left[ \prod_{j=k+1}^{M} \left( \frac{1}{2} + \frac{1}{2} \text{erf} \left( \frac{y(l) - \mu_j}{\sqrt{2}\sigma} \right) \right) \right] e^{-\frac{(y(l) - \mu_{l|l})(y(l) - \mu_{l|l})^2}{2\sigma^2}} dy(l),$$

where $w = L(W, \mathcal{P}, \text{index}(l))$, $\beta_i = d_i/d(l)$, $a_i = \text{PRR}^{-1}(\min\{\beta_i, \frac{1}{\beta_i}\})$, and the links are enumerated such that $a_1 \leq a_2 \leq \ldots \leq a_{M-1}$, with $a_0 = -\infty, a_M = \infty$ and $y_i$ is the SNR experienced by link $l_i$, and $g(\beta_i) = \begin{cases} a_i, & \text{if } \beta_i \geq 1 \\ \infty, & \text{if } \beta_i < 1 \end{cases}$

**Proof.**

From Theorem 2:

$$d_{y(l)}(y(l)|l \uparrow) = \int_{-\infty}^{\infty} y(l)^m d_{y(l)}(y(l)|l \uparrow) dy(l)$$

$$d_{y(l)}(y(l)|l \uparrow) = \frac{1}{w} \sum_{k=0}^{M-1} \int_{a_k}^{a_{k+1}} \frac{1}{\sigma \sqrt{2\pi}} \left[ \prod_{i=1}^{k} \left( \frac{1}{2} + \frac{1}{2} \text{erf} \left( \frac{g(\beta_i) - \mu_i}{\sqrt{2}\sigma} \right) \right) \right]$$

$$\left[ \prod_{j=k+1}^{M} \left( \frac{1}{2} + \frac{1}{2} \text{erf} \left( \frac{y(l) - \mu_j}{\sqrt{2}\sigma} \right) \right) \right] e^{-\frac{(y(l) - \mu_{l|l})(y(l) - \mu_{l|l})^2}{2\sigma^2}} dy(l),$$

44
Therefore,
\[
E(y(l)^m | l \uparrow) \approx \frac{1}{w} \sum_{k=0}^{M-1} \int_{a_k}^{a_{k+1}} \frac{y(l)^m}{\sigma \sqrt{2\pi}} \left[ \prod_{i=1}^{k} \left( \frac{1}{2} + \frac{1}{2} \text{erf} \left( \frac{g(\beta_i) - \mu_{y_i}}{\sqrt{2}\sigma} \right) \right) \right] \\
\left[ \prod_{j=k+1}^{M} \left( \frac{1}{2} + \frac{1}{2} \text{erf} \left( \frac{y(l) - \mu_{y_j}}{\sqrt{2}\sigma} \right) \right) \right] e^{-\frac{(y(l) - \mu_{y_j})^2}{2\sigma^2}} dy(l),
\]

3.1.3 Tuning power for matching performance

In this subsection, we present a procedure to calculate the optimal power attenuation to achieve the consistent performance repetition across two different environments. Equipped with the analytical expressions for expected network performance metrics we can define optimal power scaling rule for consistent performance across testbed and deployment environments. Consider a wireless network \( W \) with inter-node distances \( \{d_j\}_{j=1}^{m} \) and its scaled version \( \tilde{W} \), with inter-node distances \( \{\tilde{d}_j = \alpha d_j\}_{j=1}^{m} \) in RF environments characterized by log-normal scale model parameters \( (n, \sigma) \) and \( (\tilde{n}, \tilde{\sigma}) \) respectively. We assume that transmit power in the network \( \tilde{W} \) is variable through:
\[
\tilde{P}_t = \tilde{P}_0 + \beta
\]
where \( \beta \) is the power attenuation or amplification in the scaled network.

Since the scaled vector in general reduces the node distances for convenient testing, \( \beta \) in general is a negative value indicating power attenuation. As a result the scaled network realizations \( \tilde{W}(\beta) \) depends on \( \beta^1 \). The optimal power attenuation is defined as:

\[
E[f|W, \mathcal{P}] = E[f|\tilde{W}(\beta), \mathcal{P}]
\]

\(^1\)We could also introduce spatial variations in \( \beta \) across the nodes to influence the width of the resulting spectrum, for a better match.
3.1.4 Calculating the variance of performance metrics

With Theorem 4 in this chapter, we derive an equation to calculate the moments of the SNR of a link conditioned on being chosen over all other links. With Theorem 3 in chapter 2, we derive an equation to calculate an expected performance metric. In this subsection, we derive an equation to calculate the variance of performance metric (Theorem 5). The variance of performance metric $\text{Var}[f|W, \mathcal{P}]$ is computed as variance over the linkset $l_i$, weighted by the link usage spectrum $L(W, \mathcal{P}, i)$:

$$\text{Var}[f|W, \mathcal{P}] = \sum_{i=1}^{n} \text{Var}(f(y_i, d_i|l_i \uparrow))L(W, \mathcal{P}, i)$$

**Theorem 5.** The variance of performance metric, $\text{Var}(f(y_i, d_i|l_i \uparrow))$, of the link $l_i$ can be calculated as below:

$$\text{Var}(f(y_i, d_i|l_i \uparrow)) \simeq (f'(E(y_i|l_i \uparrow), d_i))^2 \text{Var}(y_i|l_i \uparrow)$$

, where $y_i$ is SNR (dB) values of link $l_i$, $d_i$ is the length of $l_i$

**Proof.** We review the straightforward argument based on the Taylor expansion of the function $f(y, d)$ with respect to its first argument.

$$\text{Var}[f(y_i|l_i \uparrow)] = \text{Var}[f(\mu_{y_i} + (y_i - \mu_{y_i})|l_i \uparrow)]$$

Using the Taylor expansion, $f(x) = \sum_{k=0}^{n} \frac{f^{(n)}(a)}{n!} (x - a)^n$ We have

$$\simeq \text{Var}[f(\mu_{y_i}|l_i \uparrow) + f'(\mu_{y_i}|l_i \uparrow)(y_i - \mu_{y_i}|l_i \uparrow)]$$

$$= (f'(E(y_i|l_i \uparrow)))^2 \text{Var}(y_i|l_i \uparrow)$$

\[\square\]
Case study: f is the number of transmissions

In this subsection, we present a case study of variance calculation when $f$ is the number of transmission, $W$ is a chain topology or 2-dimensional grid topology, and $P$ is collection tree protocol (CTP) which adopts $PRR \times D$ for routing metric. We can calculate variance of the number of transmissions using Theorem 5. However, in the last hop, we have to count an extra transmission due to the short distance to destination (edge effect).

**Corollary 1.** For a realistic topology with limited length from source to destination, and if $f$, performance metric, is the number of transmissions, and if $g$, the extra transmissions, is evenly distributed between 0 or 1 (or more), we can calculate the expected variance of the total number of transmissions as:

$$
\mu_g = \frac{n}{2} \cdot 0 + \frac{n}{2} \cdot 1 = \frac{1}{2}
$$

, where $n$ is the sample size.

$$
V(g) = \frac{n}{2} \cdot 0.25 + \frac{n}{2} \cdot 0.25 = 0.25
$$

Because the extra transmissions can be 1 or more,

$$
V(g) \geq 0.25
$$

$$
\sigma_{f+g} = \sqrt{V(f) + V(g)}
$$

$$
\sigma_{f+g} \geq \sqrt{V(f) + 0.25}
$$

### 3.1.5 Upper/lower bounds of the Protocol performance

In this subsection, we present a study on deterministic prediction of the upper bound and the lower bound of the WSN protocol performance. For a RSSI versus log of distance (10$\log_{10}(d/d_0)$), $d$ is distance, $d_0$ is unit distance) graph, $Y = -\eta X + P_0$, where $\eta$ is path loss exponent, $P_0$ is RSSI at $d_0$. We provide the upper/lower bound when $\eta$, $\sigma$, and $P_0$ are constant.

47
Upper/lower bound of WSN routing protocol performance when path loss exponent is fixed

We present the Upper/Lower bound of wireless protocol performance in this subsection when $\eta$ is fixed.

**Definition 3.** $S = (y_1, y_2, ..., y_n)$ is a set of spatially varied RSSI values of the set of links of all different lengths, $L = (l_1, l_2, ..., l_n)$

**Definition 4.** $\bar{S} = (\bar{y}_1, \bar{y}_2, ..., \bar{y}_n)$ is a set of RSSI values of the fitted (by regression process) RSSI vs. log-distance graph, $Y = -\eta X + P_0$.

Fitted red lines of Figure 1 in Question 1 represents $\bar{S}$.

**Definition 5.** $Z = (z_1, z_2, ..., z_n) = ((y_1 - \bar{y}_1), (y_2 - \bar{y}_2), ..., (y_n - \bar{y}_n))$

**Definition 6.** $z_M = \max\{z_i | 1 \leq i \leq n\}$

**Definition 7.** $z_m = \min\{z_i | 1 \leq i \leq n\}$

**Definition 8.** $E[f|W, P, P_0]$ is the expected performance $f$ with the environment $W$, protocol $P$, RSSI at $d_0 P_0$

$E[f|W, P, P_0]$ can be calculated with Theorem 3.

**Definition 9.** $P_0^*(S)$ is the projected RSSI at $d_0$ of the fitted (by regression process) graph, $Y = -\eta X + P_0^*(S)$, with the given $S = (y_1, y_2, ..., y_n)$

**Theorem 6.** When an application uses PRR×d as the metric for choosing forwarding links, and when we have a RSSI vs. log-distance graph with RSSI at $d_0 P_0$, and when the performance metric increases as $P_0$ increases, the every possible expected performance is bounded as below:

$$E[f|W, P, P_0^*(S_m)] \leq E[f|W, P, P_0^*(S)] \leq E[f|W, P, P_0^*(S_M)]$$

, where $S_m = ((\bar{y}_1 + z_m), (\bar{y}_2 + z_m), ..., (\bar{y}_n + z_m))$, $S_M = ((\bar{y}_1 + z_M), (\bar{y}_2 + z_M), ..., (\bar{y}_n + z_M))$
Proof. \[ P_0^*(S) = P_0 + P_0^*(Z) \]

\[ z_m \leq z_i \leq z_M \]

, where \( 1 \leq i \leq n \)

\[
\frac{n \cdot z_m}{n} \leq P_0^*(Z) \leq \frac{n \cdot z_M}{n}
\]

Therefore,

\[
P_0 + z_m \leq P_0^*(S) \leq P_0 + z_M
\]

\[ \equiv E[f|W, \mathcal{P}, P_0^*(S_m)] \leq E[f|W, \mathcal{P}, P_0^*(S)] \leq E[f|W, \mathcal{P}, P_0^*(S_M)] \]

\[ \square \]

Theorem 6 shows the deterministic upper bound and lower bound of the protocol performance of WSN protocols. We assume the variance of link qualities is come from the spatial RSSI variance. The upper/lower bound are calculated with the maximum difference from the fitted RSSI versus log of distance graph, \( Y = -\eta X + P_0 \). For example, if the maximum RSSI difference from the measured and fitted RSSI vs. log(distance) graph is 5 dB, then we calculate the upper bound as the expected performance with \( Y = -\eta X + P_0 + 5 \) with Theorem 3.

**Upper/lower bound of WSN routing protocol performance when \( \eta \) is bounded**

We extend Theorem 6 by supposing that PLE is bounded instead of having a fixed value. Suppose \( \eta_1 \leq \eta \leq \eta_2 \).

**Definition 10.** \( E[f|W, \mathcal{P}, \eta, P_0] \) is the expected performance \( f \) with the environment \( W \), protocol \( \mathcal{P} \), path loss exponent \( \eta \), RSSI at \( d_0 \) \( P_0 \)

\[ \eta_m = \arg\min_{\eta} \{E[f|W, \mathcal{P}, \eta, P_0]\} \]  
(3.1.3)
Theorem 7. When an application uses PRR xd as the metric for choosing forwarding links, and when we have a RSSI vs. log-distance graph with RSSI at \(d_0\ P_0\), path loss exponent \(\eta\) is bounded as \(\eta_1 \leq \eta \leq \eta_2\), and when the performance metric increases as \(P_0\) increases, the every possible expected performance is bounded as below:

\[
E[f|W, P, \eta, P_0^*(S_m)] \leq E[f|W, P, \eta, P_0^*(S)] \leq E[f|W, P, \eta, P_0^*(S_M)]
\]

, where \(S_m = ((\bar{y}_1 + z_m), (\bar{y}_2 + z_m), \ldots, (\bar{y}_n + z_m))\), \(S_M = ((\bar{y}_1 + z_M), (\bar{y}_2 + z_M), \ldots, (\bar{y}_n + z_M))\)

Proof. By extending Theorem 1,

\[
E[f|W, P, \eta_m, P_0^*(S_m)] \leq E[f|W, P, \eta, P_0^*(S)]
\]

By equation (1),

\[
E[f|W, P, \eta_m, P_0^*(S)] \leq E[f|W, P, \eta, P_0^*(S)]
\]

By equation (2),

\[
E[f|W, P, \eta, P_0^*(S)] \leq E[f|W, P, \eta_M, P_0^*(S)]
\]

By extending Theorem 1,

\[
E[f|W, P, \eta_M, P_0^*(S)] \leq E[f|W, P, \eta_M, P_0^*(S_M)]
\]

Therefore,

\[
E[f|W, P, \eta_m, P_0^*(S_m)] \leq E[f|W, P, \eta, P_0^*(S)] \leq E[f|W, P, \eta_M, P_0^*(S_M)]
\]
We present the validation study of the variance and the upper/lower bounds of the protocol performance in section 4.5.3.

3.2 Validation

In this section, we present analytical, simulation, and experimental studies of repeating protocol performance across testbed and deployed environments.

**Physical & link layer** In our experiments we use 802.15.4 transceivers embedded in popular sensor network platforms such as TelosB and MicaZ. CC2420 radio is compatible with the 2.4GHz 802.15.4 standard. 802.15.4 standard wireless physical layer employs block direct-sequence spread spectrum code with 2MChip/s chip rate and 250 kbps data rate to achieve processing and coding gain. The transmitter modulates the carrier using offset quadrature phase shift keying (O-QPSK) with half-sine shaping which is equivalent to minimum shift keying (MSK) modulation which has the following Bit Error Rate:

\[
\text{BER} = Q(\sqrt{2y/\text{PG}/\text{CG}}) = \frac{1}{2}\text{erfc}(\sqrt{y/\text{PG}/\text{CG}}) \tag{3.2.1}
\]

, where \(y\) gives the SNR. The processing gain (PG) for 802.15.4 is given by 10 log(2/0.25)=9 dB. The coding gain (CG) depends on the increased Hamming distance between the codes and is a function of the SNR itself. For low a packet error rate region coding gain can be approximated as 2 dB [22].

Thus, the packet reception rate (PRR) equation is

\[
P_{\text{RR}} = \left(1 - \frac{1}{2}\text{erfc}(\sqrt{x})\right) \times \text{packet\_size}
\]

We note that the bit-error-rate approximation given in Equation 3.2.1 assumes coherent demodulation using carrier phase information. Practical transceiver designs use non-zero IF and noncoherent demodulation. The non-ideal receiver structures
can be approximated with SNR reduction or equivalently increase in the noise floor ($P_0$) causing only a horizontal shift in the PRR curve.

We present to studies. The micro study uses a dense set of nodes in one dimension to provide granularity in the choice of link lengths. The macro study uses a coarse two dimensional topology to provide granularity in the choice of link direction towards source.

3.2.1 Micro study: chain one-dimensional topology

Experimental setup
We set up a chain topology with total 20 TelosB sensor nodes, as done in [20]. Here, we define D2 as the chain topology with the node separation of 6 ft, and D1 as the same topology with node separation of 3 ft. Each node is elevated about 4 inches from the ground. TelosB mote is equipped with CC2420 radio and USB serial for communication. The node 0 is set to be the source, and the node 19 is set to be the destination. Every two seconds, the source node produces and sends a packet. For each case, we gathers about 1,000 packets generated by the source. We logged all the paths that each packet has gone through, and only counts the body parts (i.e. excludes the first and the last hops of the paths) to calculate the average link length, because usually the first and the last hops are composed of the shorter links than the remainder links of the path. The metric used for these experiments are as follows: the average link length and the number of transmissions.

Analytical prediction of expected performance requires knowledge of the RF environmental parameters: PLE ($\eta$) of indoor and outdoor and standard deviation of RSSI (dB) ($\sigma$). We measured the RSSIs of links in a corridor in the second floor of Dreese Lab and on the top of a parking garage building. For the indoor experiment, we measured RSSIs at 20 different distances (1 ∼ 20 unit, where 1 unit = 3 ft) within
the maximum communication range with the highest transmission power (0 dBm). For the outdoor experiment, 10 measurements were taken from the distance of 1 ∼ 10 unit distances where 30 ft seems to be the maximum communication range with the same transmission power, 0 dBm. Table 3.1 presents the summary of measured RF propagation environment. We also used reported radio sensitivity of -94 dBm to adjust for the noise power $P_0$.

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Indoor</th>
<th>Outdoor</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\eta$</td>
<td>1.7555</td>
<td>2.2776</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>5.0 dB</td>
<td>4.2 dB</td>
</tr>
</tbody>
</table>

Table 3.1: Log normal model variables for indoor and outdoor RF environments

**Messaging layer and results**

We test a messaging layer protocol, CTP [6]. CTP is a tree-based collection protocol. Nodes generate routes to roots using a routing gradient. In micro study, CTP uses ETX as the routing metric. ETX implicitly favors long links over short links because each node selects the path with the minimum number of expected transmissions.

**Experimental design** TelosB mote uses 2.4 GHz frequency and provides 8 different transmission power levels: 31 (0 dBm), 27 (-1 dBm), 23 (-3 dBm), 19 (-5 dBm), 15 (-7 dBm), 11 (-10 dBm), 7 (-15 dBm), 3 (-25 dBm) [27]. We also can attach various attenuators according to attenuation levels (1 dB, 3 dB, and etc.) to modify transmitted and receive power. We used a 3 dB attenuator to construct a reasonable communication environment among 20 sensor nodes in a compact indoor area. We
tested 0 dBm for D2, -1 dBm and -6 dBm for D1 at outdoor environment. For scaling node distances by half (D2 → D1), while keeping the PLE constant (outdoor) the required transmit power attenuation is given by $-10 \times 2.2776 \times \log_{10}2 = -6.8514$ dB [29]. For indoor tests, we experimented -18 dBm at 3 ft spacing (D1) to determine the feasibility of mapping the performance across networks with different node spacings and PLEs.

**Results**

Figure 3.1: Link usage spectrum. (1) analytical with Theorem 2, (2) Monte-Carlo simulation

Figure 3.1 shows the analytical and Monte-Carlo simulation results of the link usage spectra for indoor and outdoor environments with specified transmission powers for outdoor (D2 and D1) and indoor (D1) case. We observe outdoor D2 at 0 dBm, outdoor D1 at -6 dBm and indoor D1 at -18 dBm produce similar link usage spectra.

Figure 3.2 shows the analytical average SNR values of each links chosen over all other links for each test environments. In figure 3.2, we compare the conditional mean SNR of link of length $d$ given it was chosen by the forwarding protocol over
all other links and the unconditional mean SNR of a link length $d$. We observe that $E[SNR|d_f = d] > E[SNR|d]$. 

Figure 3.3 shows the analytical average PRR values of each links chosen over all other links for each test environments. The conditional mean PRR given the event $d_f = d$, is calculated by taking the expected conditional SNRs of links that are chosen. Again we see the average PRR of the chosen links is much higher than the average PRR of all links in figure 3.3. ($E[PRR|d_f = d] > E[PRR|d]$ as discussed in Section 2).

Figure 3.4 compares the cumulative link usage spectra analytically calculated (1) and experimentally collected (2). Experimental results validate the correctness of the analytical study.

Next, we validate analytical predictions in deployment experiments using two major performance metrics, link length and end-to-end delay. Figure 3.5(1) shows the average link length taken from experiments. We can see the strong similarity between outdoor D2 (0 dBm), D1 (-6 dBm), and indoor D1 (-18 dBm) as predicted with the analytical study (details shown in table 3.2). Figure 3.5(2) shows the average
Figure 3.3: PRR of the chosen/all links. Analytical (Theorem 3)

Figure 3.4: Cumulative link usage spectrum. (1) analytical (Theorem 2), (2) experiment

end-to-end delay (the number of transmissions) taken from experiments. We also observe the strong similarity between three test cases.

Table 3.2 shows the average number of transmissions taken from experiments with the specified transmission powers. Table 3.3 shows the average link lengths taken from experiments and compares them with the analytical results. We assume the topology is infinite to calculate the expected link length analytically. However, because the topology in experiments is limited, the link lengths of marginal hops are usually shorter than those of the intermediate hops. Table 3.2 and 3.3 show strong
Figure 3.5: (1) Average number of transmissions (experimental), (2) average link length (experimental)

Table 3.2: Comparisons of performance metrics: End-to-end delay indoor and outdoor

Table 3.3: Comparisons of performance metrics: link length indoor and outdoor (experimental : calculated without the first & last hop)
similarity in the performance of three matching networks (outdoor D2 with 0 dBm, outdoor D1 with -6 dBm, and indoor D1 with -18 dBm) in experimental results. This is consistent with the analytically calculated optimal attenuation factors for matching the end-to-end delay performance from the outdoor D2 to the outdoor D1 $\beta_1 = -6.9$ dB, from the outdoor D1 to the indoor D1 $\beta_2 = -12.3$ dB and from the outdoor D2 to the indoor D1 $\beta_3 = -19.2$ dB. Overall, experimental results show little more transmissions and little shorter link length than analytically predicted. Because, the last hop to destination should be shorter than other hops (edge effect), and the first hop usually tends to be shorter than other hops, we can consider little (about 1) more transmission reasonable.

3.2.2 Macro study: grid two-dimensional topology

Experimental setup

We set up grid topologies in four major WSN testbeds (KanseiGenie, Motelab, Tutor-net, TWIST) with total 16 TelosB or TmoteSky sensor nodes. TelosB and TmoteSky motes are equipped with CC2420 radio and USB serial for communication. The node (1,1) is set to be the source, and the node (4,4) is set to be the destination. Every two seconds, the source node produces and sends a packet. For each experiment with a specific transmission power, we gathers about 1,000 packets generated by the source. We logged the path that each packet has gone through. The metric used for these experiments are as follows: the progress to destination and the number of transmission.

To perform analytical study to be compared with the experimental results, we measure PLE ($\eta$) and standard deviation of RSSI ($\sigma$) in four testbeds and present in table 3.4.
Table 3.4: Log normal model variables for RF environments of testbeds

<table>
<thead>
<tr>
<th>Metrics</th>
<th>KanseiGenie</th>
<th>Motelab</th>
<th>Tutornet</th>
<th>TWIST</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\eta$</td>
<td>1.809</td>
<td>3.995</td>
<td>4.800</td>
<td>2.697</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>5.0 dB</td>
<td>8.5 dB</td>
<td>8 dB</td>
<td>9 dB</td>
</tr>
</tbody>
</table>

Messaging layer and results

We test a messaging layer protocol, CTP. We customized CTP, which uses ETX as the default routing metric, to use $\text{PRR} \times d$ as the routing metric.

Experimental design TelosB and TmoteSky motes use 2.4 GHz frequency and provide 8 different transmission power levels. We carefully choose 4×4 (16) nodes, which form grid, in four testbeds. KanseiGenie has natural grid topology and can have the maximum 16×16 grid topology. Motelab and Tutornet do not have natural grid topology, therefore we try to choose 16 nodes to form grid-like topology. TWIST has natural grid shape topology, but walls divide nodes and inter-node distance are not the same for row and column. We test multiple transmission powers (0 dBm, -5 dBm, -7 dBm, -10 dBm, -15 dBm) to find sets of transmission powers which make CTP perform comparably with given topologies in four testbeds.

Results To validate Theorem 2 (calculates the link usage spectrum) in the two-dimensional topology, we present cumulative usage weights of links with the progress to destination (how much progress can be achieved with links). Figure 3.6(1) shows the analytical results and figure 3.6(2) shows the experimental results of KanseiGenie. Figure 3.7 shows the results of Motelab. Figure 3.8 shows the results of Tutornet. Figure 3.9 shows the results of TWIST. Overall, analytical results match experimental results in all four testbeds. However, there are gaps between analytical and experimental results for all testbeds. There can be two reasons. First, for Result ??, we
assume infinite topology by which we collect the averaged distribution of link selection. But, with testbed experiments, we test $4 \times 4$ grid topology with limited possible path length. Good long links, bad medium links, and good short links combination contributes to the specific experimental results of KanseiGenie and Motelab. Second, we consider interference from external 802.15.4 traffic, or from 802.11 traffic as the source of variance of link selection, especially for lower power cases.

Figure 3.6: Cumulative usage weights at KanseiGenie. (1) analytical, (2) experimental. transmission power = (+:0, $\times$:-5, $\Box$:-10, $\blacksquare$:-15) dBm

Figure 3.7: Cumulative usage weights at Motelab. (1) analytical, (2) experimental. transmission power = (+:0, $\times$:-5, $\Box$:-10, $\blacksquare$:-15) dBm
To validate how Theorem 3 (calculates the expected performance) performs in the two-dimensional topology, we present the analytical results of the end-to-end delay (the total number of transmissions) predicted by Theorem 3 and compared them with the experimental results in figure 3.10. Overall, analytical results match experimental results in all four testbeds. The analytical and the experimental studies on the progress to destination metric are not presented due to space limitation.
Figure 3.10: Number of transmissions. (1) analytical, (2) experimental. K: KanseiGenie, M: Motelab, TU: Tutornet, TW:TWIST

According to figure 3.10, we can find a set of transmission powers which achieve comparable performance (the number of transmissions) in four testbeds. Table 3.5 shows examples of performance matching across four testbeds. Overall, experimental results show a little more transmissions than analytically predicted. Because, the last hop to destination should be shorter than other hops (edge effect), we can consider this reasonable.

<table>
<thead>
<tr>
<th></th>
<th>KanseiGenie</th>
<th>Motelab</th>
<th>Tutornet</th>
<th>TWIST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tx. Power</td>
<td>-7dBm</td>
<td>-10dBm</td>
<td>-7dBm</td>
<td>-10dBm</td>
</tr>
<tr>
<td>Analytical</td>
<td>1.55</td>
<td>1.61</td>
<td>1.68</td>
<td>1.64</td>
</tr>
<tr>
<td>Experimental</td>
<td>2.33</td>
<td>2.09</td>
<td>2.43</td>
<td>2.04</td>
</tr>
<tr>
<td>Transmissions</td>
<td>-10dBm</td>
<td>-15dBm</td>
<td>-10dBm</td>
<td>-15dBm</td>
</tr>
<tr>
<td>Analytical</td>
<td>2.25</td>
<td>2.26</td>
<td>2.00</td>
<td>2.49</td>
</tr>
<tr>
<td>Experimental</td>
<td>3.10</td>
<td>3.17</td>
<td>3.20</td>
<td>3.06</td>
</tr>
</tbody>
</table>

Table 3.5: Total transmissions matching examples
3.3 Conclusion

In this chapter, we presented a technique for achieving repeatable performance across different networks deployed in diverse set of RF environments. Our technique accommodates network pairs whose signal propagation characteristics and inter-node spacings may be different. The method is executed in three steps: First, the link usage distribution is calculated analytically/ Second, conditional moments of the link quality are calculated for each link chosen by the forwarding protocol. Third, link usage and link quality statistics are aggregated to compute expected protocol performance using a user-specified metric for each network and to select the optimal transmission power adjustment to have same expected protocol performance across the networks.

We provide a multi testbed case study to validate the proposed method.

It is important to acknowledge the limitations of our tool, as presented. The tool targets consistency to a single chosen performance metric. Our experiments however indicate that this limitation is not particularly restrictive in practice, since attempting to achieve optimal consistency with respect to one metric seems to result in achieving good consistency with respect to other performance metrics as well. Moreover, the tool requires the forwarding protocol to be specified in order to calculate the link usage probabilities. Alternatively, link usage and conditional link SNR moments can be computed empirically from in situ observations of packet source-destination pair.
CHAPTER 4

PERFORMANCE REPEATABILITY OF WIRELESS PROTOCOLS WITH DATA DRIVEN METHOD USING RF RESOURCE SPECIFICATION

In this chapter, we study generalizing and automating the process of achieving the performance repeatability within and across environments. Specifically, we study performance repeatability of wireless protocols with data driven method using detailed RF resource specification. We designed and implemented RS-Profiler that collects RF Resource Specification data of WSN testbeds efficiently. In this chapter, first, we present accurate methods to predict the performance of WSN protocols for 1-hop and also for cumulative routing metric, and for any random topologies in WSN testbeds based on the RF resource specification data. We validate our methods by comparing with experimental results from a WSN testbed. Second, we present the analytical and experimental study on performance repeatability of WSN routing protocols in a WSN testbed over 18 different $4 \times 4$ grid wireless networks. About 94% of cases, our algorithm correctly predicts the expected performance, the number of transmissions, except the highly unstable cases due to the temporal RSSI variances. Third, we present the Monte-Carlo simulation study to predict the expected variance of protocol performance. We compared the expected variances of CTP protocol using $PRR \times d$ predicted with the analytical method presented in section 3.1.4 with the Monte-Carlo
simulation results and with the experimental results. We also present the validation study of upper/lower bound of protocol performance presented in section 3.1.5. We compared the analytical results with the experimental results.

4.1 Resource specification data collection with Profiler

Complex spatio-temporal dynamics are characteristic of WSN platforms. Depending upon platform capabilities, some of their uncertainty factors are both observable and controllable (e.g., co-channel interference may be controlled by supporting channel selection mechanisms in the experimentation infrastructure), others are observable but not controllable (e.g., slow changing large-scale path loss in static networks), and still others are neither (e.g., fast changing small-scale fast fading).

To collect observable uncertainty factors, we develop an efficient off-line strategy for estimating Resource Specification factors, e.g., wireless path loss, packet-error-rate as a function of SINR, and interference from co-existing networks. We developed Resource Specification probing application (RS-Profiler) which collects RSSI and LQI information of all links. A sender will broadcast packets with sequence number in its time slot allotted by TDMA scheduler. Therefore, we can calculate the packet reception rate for each link with the sequence numbers. RS-Profiler collects noise floor information for each node and for each channel.

With the given topology map and collected RSSI values, PLE can be calculated for each direction or for each specific set of nodes. With collected PRR values for each link, connectivity map can be processed. With collected noise floor values for each channel, interference map can also be processed. RS-Profiler collects PRR values of neighboring links, then inter-link reception correlation [33], which can be used for inferring protocol performance of opportunistic routing protocols, can be processed off-line.
4.1.1 RS-Profiler details

We developed RS-Profiler which collects RF resource specification data, which are RSSI, LQI, PRR for each link and noise floor for each node. We modified SWAT [8] for TelosB programs and implemented higher tier programs by ourselves. With RS-Profiler, we can specify the packet sender, number of packets to send, inter-packet time, transmission power, and transmission channel. We also can specify the type of RF data to collect: (RSSI, LQI, PRR) pair or noise floor.

The client program working possibly anywhere connects ports of remote servers that are connected (usually via serial communication) to sensor nodes. The client program reads a configuration file that defines the set of commands that sensor nodes will receive and execute. An example format of the configuration file is presented in Table 4.1. The first command is translated as the sensor node 398 will send 100 packets with inter-packet time 10 msec, the channel 26, and the transmission power 31 (0 dBm). The last two numbers are for the extention.

\[
\begin{array}{cccccc}
10 & 100 & 398 & 26 & 31 & 0 0 \\
10 & 100 & 395 & 26 & 31 & 0 0 \\
10 & 100 & 392 & 26 & 31 & 0 0 \\
\vdots
\end{array}
\]

Table 4.1: Example command configuration for client program

Below is the detailed list of RF environmental metrics collected directly or indirectly with RS-Profiler:
• For each node, for each transmission power, for each channel, we can get neighbors

• For each node, for each neighbor, for each transmission power, for each channel, we can collect RSSI, LQI, and PRR. From these low level data, we can get link asymmetry distribution, path loss exponent (PLE), RSSI variance, and packet delivery temporal and spatial correlation

• For each node, for each channel, we can get noise floor. From noise floor collection, we can get noise floor distribution for the whole WSN testbed

• For each hour, we can get the relation between SNR and PRR

4.1.2 Scalability problem of RS-Profiler

In case that number of packets is 100, and inter-packet time is 50 ms, for large WSN testbeds, e.g. KanseiGenie has 400 TelosB nodes, the shortest processing time can be calculated as below:

• To get (RSSI, LQI, PRR) pair, we need 4266.67 minutes (16 channels \times 8 \text{ power levels} \times 400 \text{ nodes} \times 100 \text{ packets} \times 50 \text{ ms inter-packet time} = 4266.67 \text{ min})

• To get noise floor, we need 8 minutes (16 channels \times 10 \text{ sec} = 3 \text{ min})

Therefore, total profiling time is 4266.67 min \simeq 71 \text{ hours}. With this time length, it will be very hard for any WSN testbed to collect RF environmental metrics very often like daily.
4.2 Accurate WSN routing protocol performance prediction based on RF resource specification

In this section, we present algorithms that predict the expected link selections of WSN protocols using 1-hop metric (e.g. $PRR \times D$) and cumulative metric (e.g. ETX) as routing metric. By using RS-Profiler, we have all the required inputs for algorithms.

4.2.1 Performance prediction for 1-hop metric

For 1-hop routing metric, we select $PRR \times D$ for our case study. RS-Profiler provides the expected PRR data for each directed link. For each node $v \in V[G]$, calculate $PRR(u, v) \times D$ for each neighbor u, where D is the progress to destination. We can predict the expected path of routing algorithms using $PRR \times D$ with the greedy algorithm, which is to construct the expected path selection by tracing the nodes which have the highest $PRR \times D$ value from the source to the destination. Details are presented in Algorithm 1.

**Algorithm 1 (CTP-PRRxD Algorithm).**

$CTP-PRRxD(G, PRR, D, src, dst)$

1: $u \leftarrow src$
2: $tx \leftarrow 0$
3: while ($u \neq dst$)
4: $MAX \leftarrow 0$
5: $next \leftarrow 0$
6: for each edge $(u, v) \in E[G]$ do
7: if $PRR[u, v] \cdot D[u, v] \geq MAX$ then
8: $MAX \leftarrow PRR[u, v] \cdot D[u, v]$
9: $next \leftarrow v$
4.2.2 Performance prediction for cumulative metric

For the cumulative routing metric, we select ETX for our case study. We present CTP-ETX algorithm in Algorithm 2 to calculate the expected path selection of WSN routing protocol using ETX as routing metric. RS-Profiler provides PRR data for each directed link.

**Algorithm 2 (CTP-ETX Algorithm).**

\[
\text{CTP-ETX}(G, PRR, s)
\]

1: for each vertex \( v \in V[G] \) do
2: \( d[v] \leftarrow \infty \)
3: end for
4: \( d[s] \leftarrow 0 \)
5: for \( i \leftarrow 1 \) to \(|V[G]|\) do
6: \( \text{relaxed} \leftarrow \text{False} \)
7: for each edge \( (u, v) \in E[G] \) do
8: \( \text{if } d[v] > d[u] + \frac{1}{PRR(u,v)} \frac{1}{PRR(v,u)} \text{ then} \)
9: \( d[v] \leftarrow d[u] + \frac{1}{PRR(u,v)} \frac{1}{PRR(v,u)} \)
10: relaxed ← True
11: end if
12: end for
13: if relaxed = False then
14: exit the loop
15: end if
16: end for
17: for each edge \((u, v) \in E[G]\) do
18: if \(d[v] > d[u] + \frac{1}{PRR(u,v)} \frac{1}{PRR(v,u)}\) then
19: return False
20: end if
21: end for
22: return True

4.2.3 Validation

In our previous works [20] [21], we used the following Packet Reception Rate equation:

\[
PRR(x) = \left( 1 - \frac{1}{2} \text{erfc}(\sqrt{x}) \right)^{8 \times \text{packet.size}}
\]

(4.2.1)

, where \(x\) is SNR (dB).

We measured the SNR and PRR values in KanseiGenie [2] using RS-Profiler, which we implemented by ourselves, with 100 ms of inter-packet-interval. Figure 4.1 shows the measured SNR-vs-PRR relations.
Figure 4.1: SNR vs. PRR scatter plot in KanseiGenie

Figure 4.2: Analytical SNR vs. PRR graph and experimentally adjusted SNR vs. PRR graph of KanseiGenie

Figure 4.2 shows the original analytical SNR vs. PRR graph and the experimentally adjusted graph of KanseiGenie. We use this adjusted SNR-vs-PRR relation to predict the expected PRR of links with measured SNR values in KanseiGenie.

Experimental setup

Topologies We select a 4 × 4 grid (16 TelosB nodes) in KanseiGenie. Each TelosB node is equipped with CC2420 radio. The details of the nodes of the grid is as follows: 1((1,1), source), 2(1,2), 3(1,3), 4(1,4), 5(2,1), 6(2,2), 7(2,3), 8(2,4), 9(3,1), 71
10(3,2), 11(3,3), 12(3,4), 13(4,1), 14(4,2), 15(4,3), 16((4,4), destination). For this grid, we achieved two topologies with two different transmission powers: -15 dBm, -25 dBm at the channel 11.

Figure 4.3 shows SNR values, which are experimentally measured in channel 11 with transmission power of -15 dBm, of all directed links between all pairs of nodes in the $4 \times 4$ grid.

Figure 4.4 shows the expected PRR values using the adjusted SNR vs. PRR relation, presented in Figure 4.2, based on Figure 4.3. Figure 4.5 shows the SNR values, which are experimentally measured in channel 11 with transmission power of -25 dBm, of all directed links between all pairs of nodes in the $4 \times 4$ grid. Figure 4.6 shows the expected PRR values using the adjusted SNR vs. PRR relation, presented in Figure 4.2, based on Figure 4.5. We tested a messaging layer protocol, collection tree protocol (CTP). The node (1,1) was set to be the source, and the node (4,4) was set to be the destination. Every two seconds, the source node produced and sends a packet. For each experiment with a specific transmission power, we gathered about
1,000 packets generated by the source. We logged the path that each packet has gone through.

**WSN routing protocol with one-hop routing metric:** $PRR \times D$

For this case study, we customized CTP, which uses ETX as the default routing metric, to use $PRR \times d$ as the routing metric.

Figure 4.7 shows the expected $PRR \times D$ values, which the source (node 1) has,
Figure 4.6: Expected PRRs for all pairs links of a $4 \times 4$ grid topology with channel 11, transmission power -25 dBm in KanseiGenie

![Figure 4.6](image)

Figure 4.7: Expected $PRR \times D$ values at source (node1) with channel 11, transmission power -15 dBm

![Figure 4.7](image)

Based on the RF measurement data presented in Figure 4.3, 4.4. According to Figure 4.7 and CTP-PRRxD algorithm (algorithm 1), node 1 will choose node 12, with the coordinate (4,3), as the next hop, and the temporal RSSI variation can give chances node 15 (4,3), 8 (4,2), 11 (3,3), 14 (4,2), 4 (1,4), 7 (2,3), 13 (4,1) become the next hop of the node 1.

Table 4.2 shows the analytical and experimental results of path selection and of the performance metric with CTP with $PRR \times D$. We could predict the path

74
Table 4.2: Path selection and the performance metric with $PRR \times D$: channel 11, transmission power -15 dBm at KanseiGenie

<table>
<thead>
<tr>
<th>Path to destination</th>
<th># of transmissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytical</td>
<td></td>
</tr>
<tr>
<td>1 → 12 → 16</td>
<td>2</td>
</tr>
<tr>
<td>Experimental</td>
<td></td>
</tr>
<tr>
<td>1 → 12 → 16</td>
<td>2</td>
</tr>
<tr>
<td>1 → 11 → 16</td>
<td>2</td>
</tr>
<tr>
<td>1 → 4 → 16</td>
<td>2</td>
</tr>
<tr>
<td>1 → 14 → 16</td>
<td>2</td>
</tr>
<tr>
<td>1 → 15 → 16</td>
<td>2</td>
</tr>
</tbody>
</table>

selection and the number of transmission numbers relatively accurately. As expected, due to the temporal RSSI variations, there are multiple different path selections in experiments.

Figure 4.8: Expected $PRR \times D$ values at the source (node1) with channel 11, transmission power -25 dBm
Figure 4.9: Expected $PRR \times D$ values at node 4 with channel 11, transmission power -25 dBm

Figure 4.8 shows the expected $PRR \times D$ values, which the source (node 1) has, based on the RF measurement data presented in Figure 4.5, 4.6. Figure 4.9 shows the expected $PRR \times D$ values that node 4 has. According to Figure 4.8 and CTP-PRRxD algorithm (algorithm 1), node 1 will choose node 4, with the coordinate (1,4), as the next hop, and node 4 will choose node 12 as the next hop to the destination. Table 4.3 shows the analytical and experimental results of path selection and of the

<table>
<thead>
<tr>
<th>Path to destination</th>
<th># of transmissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytical</td>
<td>$1 \rightarrow 4 \rightarrow 12 \rightarrow 16$</td>
</tr>
<tr>
<td>Experimental</td>
<td>$1 \rightarrow 4 \rightarrow 12 \rightarrow 16$</td>
</tr>
</tbody>
</table>

Table 4.3: Path selection and the performance metric with $PRR \times D$: channel 11, transmission power -25 dBm at KanseiGenie

performance metric with CTP with $PRR \times D$. We could predict the path selection and the number of transmission numbers accurately.
WSN routing protocol with cumulative routing metric: ETX

For this case study, we used CTP, which uses ETX as the default routing metric.

Figure 4.10: Expected ETX values for all pairs links of a $4 \times 4$ grid topology with channel 11, transmission power -15 dBm in KanseiGenie

Figure 4.10 shows the expected ETX ($= \frac{1}{\text{PRR}(u,v)} \frac{1}{\text{PRR}(v,u)}$) values, which the node u has for every neighbor v, based on the RF measurement data presented in Figure 4.3, 4.4.

Figure 4.11: Expected ETX values at the source (node1) with channel 11, transmission power -15 dBm

77
Figure 4.11 shows the expected ETX values, which the source (node 1) has, based on the RF measurement data presented in Figure 4.3, 4.4.

According to figure 4.10, 4.11 and CTP-ETX algorithm (algorithm 2), node 1 will choose node 11, with the coordinate (3,3), as the next hop, and the temporal RSSI variation can give chances node 3 (1,3), 4 (1,4), 5 (2,1), 6 (2,2), 7 (2,3), 8 (4,2), 9 (2,1), 12 (3,4), 13 (4,1), 14 (4,2), 15 (4,3) become the next hop of the node 1.

<table>
<thead>
<tr>
<th>Path to destination</th>
<th># of transmissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytical</td>
<td></td>
</tr>
<tr>
<td>1 → 11 → 16</td>
<td>2</td>
</tr>
<tr>
<td>Experimental</td>
<td></td>
</tr>
<tr>
<td>1 → 4 → 16</td>
<td>2</td>
</tr>
<tr>
<td>1 → 14 → 16</td>
<td>2</td>
</tr>
<tr>
<td>1 → 11 → 16</td>
<td>2</td>
</tr>
<tr>
<td>1 → 5 → 16</td>
<td>2</td>
</tr>
<tr>
<td>1 → 12 → 16</td>
<td>2</td>
</tr>
<tr>
<td>1 → 15 → 16</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 4.4: Path selection and the performance metric with ETX: channel 11, transmission power -15 dBm at KanseiGenie

Table 4.4 shows the analytical and experimental results of path selection and of the performance metric with CTP with ETX. We could predict the path selection and the number of transmission numbers relatively accurately. As expected, due to the temporal RSSI variations, there are multiple different path selections in experiments.

Figure 4.12 shows the expected ETX \( \left( \frac{1}{\text{PRR}(u,v)} \cdot \frac{1}{\text{PRR}(v,u)} \right) \) values, which the node u has for every neighbor v, based on the RF measurement data presented in Figure 4.5, 4.6.
Figure 4.12: Expected ETX values for all pairs links of a $4 \times 4$ grid topology with channel 11, transmission power -25 dBm in KanseiGenie.

Figure 4.13: Expected ETX values at the source (node1) with channel 11, transmission power -25 dBm.

Figure 4.13 shows the expected ETX values, which the source (node 1) has, based on the RF measurement data presented in Figure 4.5, 4.6.

According to Figure 4.12, 4.13 and CTP-ETX algorithm (algorithm 2), node 1 will choose node 4, with the coordinate (1,4), then node 4 will choose node 8 to get to destination. Table 4.5 shows the analytical and experimental results of path selection and of the performance metric with CTP with ETX. We could predict the path selection and the number of transmission numbers.
4.3 Performance repeatability of WSN routing protocols in WSN testbeds

In this section, we present a study on performance repeatability of WSN routing protocols within WSN testbeds. In section 4.2.3, we presented our experimental results to validate that WSN routing protocol performance can be predicted accurately with RF resource specification. There, we tested CTP with $PRR \times D$ metric in a $4 \times 4$ grid topology at channel 11 and with transmission power of -25 dBm. Here, we try to repeat the similar performance (the number of transmissions) in any different $4 \times 4$ grid networks within KanseiGenie WSN Testbed. With the RF resource specification data collected with RS-Profiler and the performance prediction algorithm for 1-hop routing metric (presented in section 4.2.1), we execute the analytical study. Then, we execute the experimental study with multiple (17) candidates as shown in Table 4.6 to check the performance repeatability of WSN protocols.

With the original network, we have the total transmissions of 4.6 with 3 hops. With the analytical study, we predict that the candidate 5 network will have the most similar performance to the original network among the ten candidates, and we validate that with the experimental studies.

In Table 4.6, we can see that most of the cases, our analytical study accurately predict the expected routing path, the total number of transmissions, and the number
<table>
<thead>
<tr>
<th>4 × 4 grid</th>
<th>Path</th>
<th># Tx.</th>
<th># Hops</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>E</td>
<td>A</td>
</tr>
<tr>
<td>Original</td>
<td>1 → 4 → 12 → 16</td>
<td>1 → 4 → 12 → 16</td>
<td>4.3</td>
</tr>
<tr>
<td>Candidate 1</td>
<td>1 → 12 → 16</td>
<td>1 → 12 → 16</td>
<td>2.22</td>
</tr>
<tr>
<td>Candidate 2</td>
<td>1 → 7 → 16</td>
<td>1 → 7 → 16</td>
<td>2.1</td>
</tr>
<tr>
<td>Candidate 3</td>
<td>1 → 12 → 16</td>
<td>1 → 12 → 16</td>
<td>2.4</td>
</tr>
<tr>
<td>Candidate 4</td>
<td>1 → 6 → 8 → 16</td>
<td>1 → 6 → 8 → 16</td>
<td>3</td>
</tr>
<tr>
<td>Candidate 5</td>
<td>1 → 2 → 11 → 15 → 16</td>
<td>1 → 2 → 11 → 15 → 16</td>
<td>4.1</td>
</tr>
<tr>
<td>Candidate 6</td>
<td>1 → 15 → 16</td>
<td>1 → 15 → 16</td>
<td>3</td>
</tr>
<tr>
<td>Candidate 7</td>
<td>1 → 9 → 16</td>
<td>1 → 3 → 8 → 16</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 → 9 → 16</td>
<td></td>
</tr>
<tr>
<td>Candidate 8</td>
<td>1 → 14 → 15 → 16</td>
<td>1 → 14 → 15 → 16</td>
<td>3.4</td>
</tr>
<tr>
<td>Candidate 9</td>
<td>1 → 14 → 16</td>
<td>1 → 14 → 16</td>
<td>2.2</td>
</tr>
<tr>
<td>Candidate 10</td>
<td>1 → 10 → 16</td>
<td>1 → 10 → 16</td>
<td>2</td>
</tr>
<tr>
<td>Candidate 11</td>
<td>1 → 5 → 15 → 16</td>
<td>1 → 12 → 16</td>
<td>3.1</td>
</tr>
<tr>
<td>Candidate 12</td>
<td>1 → 10 → 15 → 16</td>
<td>1 → 10 → 15 → 16</td>
<td>3.3</td>
</tr>
<tr>
<td>Candidate 13</td>
<td>1 → 14 → 16</td>
<td>1 → 12 → 16</td>
<td>2.2</td>
</tr>
<tr>
<td>Candidate 14</td>
<td>1 → 8 → 16</td>
<td>1 → 12 → 16</td>
<td>2.4</td>
</tr>
<tr>
<td>Candidate 15</td>
<td>1 → 4 → 16</td>
<td>1 → 7 → 16</td>
<td>2.2</td>
</tr>
<tr>
<td>Candidate 16</td>
<td>1 → 9 → 14 → 16</td>
<td>1 → 9 → 14 → 16</td>
<td>3</td>
</tr>
<tr>
<td>Candidate 17</td>
<td>1 → 12 → 16</td>
<td>1 → 12 → 16</td>
<td>2.2</td>
</tr>
</tbody>
</table>

Table 4.6: Analytically predicted and experimented performance repetition of CTP with $PRR \times D$: channel 11, transmission power of -25 dBm at KanseiGenie WSN testbed. A: analytical, E: experimental
of hops. However, in some cases, the experimental results deviate from the analytical results. In candidate 7 case, multiple routing paths are chosen. In candidate 8 case, there is large variation in the total number of transmissions, though the routing path and the number of hops are correctly predicted. All these errors are due to the temporal RSSI variation, which is natural and unavoidable, or due to SNR measurement errors.

Table 4.7 shows the statistics of the accuracy of the performance prediction algorithm 1. Algorithm 1 correctly predicts the number of transmissions in 94% of cases.

<table>
<thead>
<tr>
<th>Correct Predictions</th>
<th>Ratio(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Hops</td>
<td>17 / 18 (94.44)</td>
</tr>
<tr>
<td>Path</td>
<td>14 / 18 (77.78)</td>
</tr>
<tr>
<td>Number of Transmissions</td>
<td>15 / 18 (83.33)</td>
</tr>
<tr>
<td>( e \leq 0.5 )</td>
<td>15 / 16 (93.75)*</td>
</tr>
</tbody>
</table>

Table 4.7: Statistics of WSN routing performance prediction. * excludes cases with large variance due to temporal RSSI variance (candidate 8, 11). \( e \) is the difference between the number of transmissions analytically predicted and experimentally measured.

### 4.4 Temporal Link Quality Variance Study

In this section, we present the temporal link quality variance study. We present the SNR variances of links in long term (about one week), and short term (about 10 seconds). We need to understand the pattern of link quality variance to predict
how often we have to measure link qualities to maintain the validity of the performance prediction and performance repetition method. And, the very unreliable experimental results with candidate 8, 11 network motivate this study to find the reason. Figure 4.14 shows the long term (one week) temporal SNR variance of total

![Figure 4.14: Long term temporal SNR variance of links of Candidate 8 (1), Candidate 10 (2) at KanseiGenie channel 11](image)

5 links. Three links belong to candidate 8 with highly variable performance, and the other two links belong to candidate 10 with very stable performance. Surprisingly, all five links show very stable SNRs varying within 2 dB from the average SNRs of links in a whole week. Table 4.11 shows the summary of Figure 4.14. Table 4.9 shows the variances of the expected number of transmissions of CTP with $P_{RR} \times d$ run on candidate 8 and 10 grid networks with the link qualities from table 4.11. Compared to the experimental results in table 4.6, CTP performed at candidate 8 network with 3 $\sim$ 9 total transmissions, and table 4.9 predicts that CTP will perform at candidate 8 networks usually 4 $\sim$ 8 total transmissions. And, for candidate 10, CTP performed with 2 total transmissions, and table 4.9 predicts that CTP will perform usually about 2 total transmissions. Therefore, using the long-term temporal link quality
variation of links of the target network, we can roughly predict the variance and the upper/lower bound of the wireless protocol performances.

We also measured short term (10 seconds) SNR variances of links. 100 packets are sent with 100 msec of the inter packet time for each 5 links. All SNR values of links varies within 1 dB from the average in Figure 4.15. Then, what is the real reason of the high variance of the performance (the number of transmissions) of candidate 8? As we can see in Figure 4.14 (1), in many cases the link qualities of links locate below 5 dB, which can be translated to 50 % of PRR in Figure 4.2. With just one dB drop, the PRR of links will precipitate to below 20 %. Therefore, by combining Figure 4.14 and 4.15, SNR long term variation make SNRs of links of candidate 8 below 5 dB, then SNR short term variation with just 1 dB variance make a lot of variation on the performance.

<table>
<thead>
<tr>
<th></th>
<th>C 8</th>
<th>C 8</th>
<th>C 8</th>
<th>C 10</th>
<th>C 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 → 14</td>
<td>5.6</td>
<td>6.9</td>
<td>6.8</td>
<td>10</td>
<td>15.4</td>
</tr>
<tr>
<td>14 → 15</td>
<td>3.6</td>
<td>4</td>
<td>4.84</td>
<td>5.73</td>
<td>12.4</td>
</tr>
<tr>
<td>15 → 16</td>
<td>4.88</td>
<td>5.37</td>
<td>6.07</td>
<td>7.83</td>
<td>14.34</td>
</tr>
<tr>
<td>1 → 10</td>
<td>0.48</td>
<td>0.76</td>
<td>0.57</td>
<td>0.85</td>
<td>0.8</td>
</tr>
<tr>
<td>10 → 16</td>
<td>0.48</td>
<td>0.76</td>
<td>0.57</td>
<td>0.85</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Table 4.8: Max, Min, Mean and Variance of SNR (dB) of links of KanseiGenie testbed channel 11 with -25 dBm transmission power measured over 180 hours period. C: Candidate 4 × 4 grid network
Table 4.9: Expected number of transmissions at Max, Min, $\mu - \sigma$, $\mu$, and $\mu + \sigma$ of SNR (dB) of links of KanseiGenie testbed Channel 11 with -25 dBm transmission power measured over 180 hours period. C: Candidate 4 × 4 grid network

<table>
<thead>
<tr>
<th></th>
<th>min</th>
<th>$\mu - \sigma$</th>
<th>$\mu$</th>
<th>$\mu + \sigma$</th>
<th>MAX</th>
</tr>
</thead>
<tbody>
<tr>
<td>C 8</td>
<td>16.84</td>
<td>7.4994</td>
<td>5.013</td>
<td>3.9475</td>
<td>3.6102</td>
</tr>
<tr>
<td>C 10</td>
<td>2.38</td>
<td>2.0658</td>
<td>2.0162</td>
<td>2.0031</td>
<td>2.0001</td>
</tr>
</tbody>
</table>

Figure 4.15: Short term temporal RSSI variance of links of Candidate 8 (1), Candidate 10 (2) at KanseiGenie channel 11

4.5 Predicting performance variance with resource specification

In section 4.2, we present algorithms that predict the expected link selections of WSN protocols using 1-hop metric (e.g. $PRR \times D$) (Algorithm 1) and cumulative metric (e.g. ETX [15]) (Algorithm 2) as routing metric. By using RS-Profiler, we have all the required inputs for the algorithms.
4.5.1 Monte-Carlo simulation study with performance prediction

Table 4.7 shows the statistics of the accuracy of the performance prediction algorithm. Algorithm 1 correctly predicts the number of transmissions in 94% of cases. We exploit the high accuracy of the performance prediction algorithm by integrating it into Monte-Carlo simulation. We generate SNR values for 240 pairs of directed links between $4 \times 4$ (16) nodes with random number generator to follow the Gaussian distribution using the RF metrics presented in Table 4.10 and Figure 4.16. Then, we apply the performance prediction algorithm for $PRR \times D$ routing metric (Algorithm 1) to predict the expected path and the expected number of transmissions. With the assigned SNR values of links, we can calculate the predicted packet reception rate (PRR) values using the graph shown in Figure 4.2, then we can calculate the expected number of transmissions for the link with $\frac{1}{PRR(u,v)}$ value, where $u$ is the sender, $v$ is the receiver.

We repeated this process 1,000 times, then calculated the mean and the variance. This simulation results will be the upper bound of the real experimental results, because in the real experiments, we may not be able to collect data for a network with bad links along the chosen path.

4.5.2 Validation of performance variance prediction methods

In this subsection, we present an analytical result that use Theorem 3, 5, Colollary 1, and Monte-Carlo simulation results that use RF resource specification and performance prediction algorithm presented in Section 4.2, and experimental results of measuring variance of performance metrics of wireless protocols.
Experimental setup

**Topologies** We select 18 4x4 grids (16 TelosB nodes) with 6 ft inter-node spacing, and union of them covers the whole KanseiGenie WSN testbed. Each TelosB node is equipped with CC2420 radio. The details of the nodes of the grid is as follows: 1((1,1), source), 2(1,2), 3(1,3), 4(1,4), 5(2,1), 6(2,2), 7(2,3), 8(2,4), 9(3,1), 10(3,2), 11(3,3), 12(3,4), 13(4,1), 14(4,2), 15(4,3), 16((4,4), destination). For these grids, we configure wireless networks with transmission power of -25 dBm at the channel 11. TelosB mote uses 2.4 GHz frequency and provides 8 different transmission power levels: 31 (0 dBm), 27 (-1 dBm), 23 (-3 dBm), 19 (-5 dBm), 15 (-7 dBm), 11 (-10 dBm), 7 (-15 dBm), 3 (-25 dBm) [27]. We chose the channel 11, because it is the least interfered channel in KanseiGenie according to our noise floor measurement study and we want to minimize the impact of the temporal RSSI variance on the variance of the protocol performance. The noise floor measurement results from a TelosB node are shown in Figure 5.3. We sampled noise floors at 100 Hz for 30 seconds for each channel.

We chose -25 dBm for the transmission power to achieve multihop topologies in 4 x 4 grids. With higher possible transmission power levels, all grids usually have 1-hop topologies in channel 11 of KanseiGenie.

<table>
<thead>
<tr>
<th>Metrics</th>
<th>KanseiGenie Channel 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Path Loss Exponent</td>
<td>2.055</td>
</tr>
<tr>
<td>RSSI Standard Dev.</td>
<td>4.93 dB</td>
</tr>
</tbody>
</table>

Table 4.10: Log normal model variables for KanseiGenie channel 11

The basic radio communication metrics of KanseiGenie channel 11 are shown in...
Table 4.10 and Figure 4.16. We applied these metrics into Theorem 3, 5 in the analytical study.

**Messaging layer** We experiment CTP as the messaging layer protocol. In experiments, CTP that originally adopts ETX as the routing metric uses $PRR \times D$. For $4 \times 4$ grids, the node at coordinate $(1,1)$ is set to be the source, and the node at $(4,4)$ is set to be the destination. Every two seconds, the source node produces and sends a packet. For each experiment with a specific transmission power, we gathers about 1,000 packets generated by the source. We logged the path that each packet has gone through.

### 4.5.3 Validation of protocol variance prediction

**Analytical results**

The link usage spectrum (Theorem 1, 2) is validated through the analytical and experimental studies in chapter 2 and in [20]. The expected m-th moment (Theorem 4)
and the expected performance of WSN protocols (Theorem 3) are validated through analytical, simulation and experimental studies in chapter 2 and in [21].

The analytical and experimental results of the variance of the WSN protocol performance (Theorem 5) is presented in Table 4.11.

### Comparison

Table 4.11 shows the analytical results of the mean and variance of the WSN protocol performance (Theorem 5, Colollary 1) with the $4 \times 4$ grid network setup presented in Section 4.5.2. The next row shows the summary (mean and variance) of the Monte-Carlo simulation results exploiting the performance prediction algorithm (Algorithm 1) with the same $4 \times 4$ grid network setup. The last row show the summary of the experimental results with the same experimental setup. In Table 4.11, the analytical

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Stdev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytical</td>
<td>3.06</td>
<td>0.6169</td>
</tr>
<tr>
<td>Monte-Carlo Simulation with Performance Prediction</td>
<td>3.15</td>
<td>0.93</td>
</tr>
<tr>
<td>Experimental</td>
<td>2.67</td>
<td>0.75</td>
</tr>
</tbody>
</table>

Table 4.11: Means and variances of performance (number of transmissions) of CTP with $PRR \times D$ analyzed/experimented over $4 \times 4$ grids in channel 11 with -25 dBm transmission power.

study and the Monte-Carlo simulation with the performance prediction algorithm for 1-hop routing metric can precisely predict the expected performance (the number of transmissions) and the variance compared to the real experiments. As we predicted
in Section 4.5.1, analytical and simulation results are the upper bounds of the experimental results, because in the real experiments, we may not be able to collect data for a network with bad links along the chosen path.

**Validation of upper/lower bound of protocol performance prediction**

Theorem 7 is the extended version of Theorem 6, and the upper/lower bounds predicted by Theorem 7 is broader than those by Theorem 6, because Theorem 7 considers the bounded PLE which broadens the upper/lower bounds. We use the experimental results presented in Table 4.6, 4.11 for validation of Theorem 6, 7. We use the RF environmental metrics presented in Table 4.10, and use Theorem 1, 2, 4, 3 to calculate the expected performance of CTP with $PRR \times D$ metric tested over $4 \times 4$ grid network, and use Theorem 6 to calculate the upper and lower bounds.

<table>
<thead>
<tr>
<th></th>
<th>Analytical with spatial RSSI variance</th>
<th>Experimental</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Upper Bound</td>
<td>Lower Bound</td>
</tr>
<tr>
<td># of transmissions</td>
<td>8.17</td>
<td>1.54</td>
</tr>
</tbody>
</table>

Table 4.12: Upper/lower of performance (number of transmissions) of CTP with $PRR \times D$ analyzed/experimented over $4 \times 4$ grids in channel 11 with -25 dBm transmission power.

The results are shown in Table 4.12. Clearly, the upper/lower bounds predicted by Theorem 6 work as upper/lower bounds of the experimented maximum and minimum transmission numbers. We can consider that with experiments, the grid networks with the bad communication situation (e.g. RSSIs are lower than others) may not produce results. Therefore, the experimented maximum number of transmissions may be lower.
than the analytically predicted. Even though we consider this point, we need tighter upper/lower bounds than predicted with Theorem 6.

4.6 Conclusion

In this chapter, we study the performance repeatability of WSN protocols as a case study of applying Resource Specification data to serve the experimentors request that as experimenters specify the RSpec values of the target experimentation infrastructure, WSN testbed finds the comparable topology. We achieve the performance repeatability through an analytical method that can accurately predict the performance of wireless routing protocols, which use 1-hop metric and cumulative metric, based on the measurement of RF resource specification. Our contributions in this chapter are as follows. First, We provide analytical method to calculate the expected performance and the path selection of wireless routing protocol based on the measured RF resource specification very accurately. Especially, our methods can be applied to any topology, not limited to chain or grid, and any size of topology. Second, we present the analytical and experimental study on performance repeatability of WSN routing protocols in a WSN testbed. Because collecting RF data of links may not be feasible in outdoor real deployments, our works will contribute in protocol performance prediction in WSN testbed and performance repetition within and across WSN testbeds. In this chapter, we presented the Monte-Carlo simulation study to predict the expected variance of protocol performance. We compared the expected variances of CTP protocol using $P_{RR} \times d$ predicted with the analytical method (theorem 5) with the Monte-Carlo simulation results and with the experimental results. We also present the validation study of upper/lower bound of protocol performance presented in section 3.1.5. We compared the analytical results with the experimental results.
CHAPTER 5
SCALABILITY OF RESOURCE SPECIFICATION PROFILING ON LARGE SCALE WSN TESTBDS

In this chapter, we study methods to predict link qualities (e.g. RSSI, SNR, PRR) without whole direct measurements. We study the scalability problem of RS-Profiler on large scale WSN testbeds, and the corresponding solutions in detail. We study how to reduce the profiling time by adopting accurate prediction on link qualities (RSSI, SNR, PRR) with partial measurements.

5.1 Complexity analysis of resource specification profiling time

In this section, we present an equation of the expected profiling time with required parameters, and show the probable actual profiling time by plugging some actual probable parameter values into the equation. Let’s define the number of nodes as N, the number of channels as C, and the number of transmission power levels as P. Let’s define the processing time of RSSI measurement for a node as $T_{\text{rssi}}$, and the processing time of noise floor measurement for a node as $T_{nf}$. Then, the total profiling time ($T(N)$) of resource specification data will be calculated as the sum of
the total RSSI measurement time of all nodes for all channels, for all transmission power levels, and the total noise floor measurement time of all nodes for all channels.

\[ T(N) = N \cdot C \cdot P \cdot T_{\text{rssi}} + C \cdot T_{\text{nf}} \] (5.1.1)

Though the complexity of profiling time on \( N \) is \( O(N) \), we cannot ignore the constant values (\( C \) and \( P \)), because those impact so much on the possible frequency of profiling process, which again will impact on the accuracy of link quality predictions (RSSI, SNR, PRR) and ultimately on the accuracy of the expected performance of the target network. To have more concrete idea on how frequently the profiling process can be executed, we present examples of profiling time by plugging very probable values into the total profiling time equation. If we plug in practical values for \( N \) with 400 (400 TelosBs in KanseiGenie) into equation 5.1.1, for \( C \) with 16 (16 CC2420 channels: ch11 ∼ ch26), for \( P \) with 8 (8 CC2420 transmission power levels: 0 dBm, -1 dBm, -3 dBm, -5 dBm, -7 dBm, -10 dBm, -15 dBm, -25 dBm), for \( T_{\text{rssi}} \) with 5 sec (100 packets with 50 msec inter-packet-interval), and for \( T_{\text{nf}} \) with 10 sec (20K samples with2 KHz sampling rate), then \( T(400) = 400 \cdot 16 \cdot 8 \cdot 5 + 1 \cdot 16 \cdot 10 \) seconds equal about 71 hours, which is impractically long time to run profiler in WSN testbeds.

If the RSSI values of a link with various transmission power levels can be predicted with a RSSI value with the highest transmission power level, then the profiling time will be reduced as follows:

\[ T(N) = N \cdot C \cdot T_{\text{rssi}} + C \cdot T_{\text{nf}} \] (5.1.2)

The complexity of profiling time on \( N \) is still \( O(N) \). If we plug in values similarly for \( N \) with 400, for \( C \) with 16, for \( P \) with 8, for \( T_{\text{rssi}} \) with 5 sec, and for \( T_{\text{nf}} \) with 10 sec into equation 5.1.2, then \( T(400) = 400 \cdot 16 \cdot 5 + 1 \cdot 16 \cdot 10 \) seconds equal about 9 hours, which is still quite long time to run profiler in WSN testbeds, but it is feasible.
If the RSSI values of links can be predicted with a noise floor measurement, then the profiling time can be reduced drastically as follows:

\[ T(N) = C \cdot T_{nf} \quad (5.1.3) \]

The complexity of profiling time on N becomes \( O(1) \). Because noise floor measurement can be done in parallel, the scaling factor N vanishes. If we plug in values similarly for N with 400, for C with 16, for P with 8, for \( T_{\text{rsi}} \) with 5 sec, and for \( T_{nf} \) with 10 sec into equation 5.1.3, then \( T(400) = 16 \cdot 10 \text{ seconds} \) equal about 3 minutes, which can be done nearly as on-demand, online profiling. In the following sections, we explore RSSI prediction methods to reduce resource specification profiling time.

5.2 Accurate prediction of RSSI values on different transmission power levels

5.2.1 Basic method

The log-normal shadowing model, a large scale fading model employed commonly in indoor and outdoor link studies, describes the received signal strength as:

\[ R(d) = P_t - PL(d_0) - 10\eta \log(d/d_0) + N_\sigma \quad (5.2.1) \]

where \( \eta \) is the path loss exponent, \( P_t \) is the transmission power, and \( PL(d_0) \) is the path loss observed at distance \( d_0 \) in dB and \( N_\sigma \) is a zero-mean Gaussian random variable with standard deviation \( \sigma \), representing spatial variations in the RF environment. The SNR at the receiver \( y(d) \) is given by the received signal power \( R(d) \) reduced by the noise power \( P_0 \):

\[ y(d) = R(d) - P_0 \text{ (in dB)} \quad (5.2.2) \]
Therefore, we can predict the RSSI of a link with the transmission power of $P_t - \alpha$ will be

$$R(d) = P_t - \alpha - PL(d_0) - 10\eta \log(d/d_0) + N_o$$

5.2.2 Refined method

In this subsection, upon this basic RSSI prediction method, we refine the RSSI prediction with a RSSI calibration method. Chen et.al. [11] argued and proved that CC2420 radio reports RSSI values erroneously when receives packets. They provided a calibrating method and a mapping table between the raw RSSI (dBm) and the calibrated RSSI (dBm). In [11], they constructed the mapping table by comparing the RSSIs reported by CC2420 radio with packet receptions and the received signal strength (RSS) (dBm) of the received packets measured with a spectrum analyzer.

In our refined RSSI prediction method, first, we send packets with the highest transmission power levels (0 dBm), then collect raw RSSI values in each receiver nodes, then calibrate the RSSI values with the mapping table from [11]. With the calibrated RSSI values, we predict the calibrated RSSI values of packets transmitted with 7 different lower power levels. The results are shown in figure 5.1.

5.2.3 Validation

In this subsection, we validate the basic and the refined RSSI prediction methods presented in section 5.2.1, 5.2.2 with RSSI measurement experiments.

In figure 5.1, we show the cumulative distributions of the prediction errors between the measured RSSI values with transmission powers of -1 dBm (1), -3 dBm (2), -5 dBm (3), -7 dBm (4), -10 dBm (5), -15 dBm (6), -25 dBm (7) and the predicted RSSI
Figure 5.1: Cumulative prediction errors based on raw RSSI and calibrated RSSI: (1) 0dB vs. -1dB, (2) 0dB vs. -3dB, (3) 0dB vs. -5dB, (4) 0dB vs. -7dB, (5) 0dB vs. -10dB, (6) 0dB vs. -15dB, (7) 0dB vs. -25dB
Figure 5.2: Prediction errors based on calibrated RSSI: (1) 0dB vs. -1dB, (2) 0dB vs. -3dB, (3) 0dB vs. -5dB, (4) 0dB vs. -7dB, (5) 0dB vs. -10dB, (6) 0dB vs. -15dB, (7) 0dB vs. -25dB
values with RSSI values measured with the transmission power of 0 dBm together with the prediction method presented in section 5.2. The target network is the original $4 \times 4$ grid in Table 4.6. We examine total 132 available directed links.

In Figure 5.1 (6), about 10% cases (0 error (dB) of x-axis) of links are not connected as predicted with the prediction method. For more than 90% of the links, RSSIs are accurately predicted within the error of 2 dB. Considering the possible 2 dB of long term temporal link quality variation, presented in section 4.4, the prediction method is quite accurate. In Figure 5.1 (7), about 50% cases (0 error (dB) of x-axis) of links are not connected as predicted with the prediction method. For more than 90% of the links, RSSIs are accurately predicted within the error of 2 dB. In figure 5.1, we also show the cumulative distributions of the prediction errors between the real calibrated RSSI values with transmission powers of -1 dBm (1), -3 dBm (2), -5 dBm (3), -7 dBm (4), -10 dBm (5), -15 dBm (6), -25 dBm (7) and the predicted calibrated RSSI values with the calibrated RSSI values measured and calibrated with the transmission power of 0 dBm together with the refined RSSI prediction method presented in section 5.2. Figure 5.1 shows the scattered plot of the prediction errors between the measured calibrated RSSI values with transmission powers of -1 dBm (1), -3 dBm (2), -5 dBm (3), -7 dBm (4), -10 dBm (5), -15 dBm (6), -25 dBm (7) and the predicted calibrated RSSI values with RSSI values measured with the transmission power of 0 dBm together with the prediction method presented in section 5.2.

As we can see, for transmission power of -1 dBm, -3 dBm, and -25 dBm, the RSSI prediction accuracy with the refined method is very similar to that with the basic method. Both methods can predict the RSSI values for more than 80% of the links within 1 dB. For -1 dBm and -3 dBm transmission power, there is not so much room for errors, and for -25 dBm, more than 50% of the links are disconnected and considered as 0 dB error by default. However, for the other 5 middle range
transmission power levels, the RSSI reporting errors by CC2420 contribute 5~10% of differences within 1 dB accuracy between the basic and the refined RSSI prediction method. By adopting the RSSI calibration, we can predict RSSI values of about 80% or more links within 1 dB accuracy with just one transmission power level RSSI measurement study. Because 1 dB in the transition region of PRR can be 50% or more difference, maintaining the RSSI prediction error within 1 dB instead of 2 dB can make huge difference in the protocol performance prediction.

5.3 Accurate prediction of RSSI with noise floor measurement

In section 2.4.4, we found that the interference level makes impact on the quality of links. Figure 5.3 shows the average noise floor values and the average interference values of all channels of KanseiGenie.

5.3.1 Prediction method

![Average noise floor and average interference values](Figure 5.3)

Figure 5.3: Noise floor and interference measurements at a node of KanseiGenie
At first we measure noise floor of KanseiGenie in three different channels: 26, 11, 12. Channel 26 is the default channel of 802.15.4 radio, and free from CTI. Channel 11 cover boundary of 802.11b channel 1. Channel 12 cover the middle part of 802.11b channel 1. Therefore, if no 802.15.4 communication is running, interference level will be: $\text{ch26} < \text{ch11} < \text{ch12}$. Figure 5.4 shows noise floor measurements at KanseiGenie of channel 26, 11, and 12. The results are quite different from our expectation: $\text{ch11} < \text{ch11} < \text{ch26}$. Even though we chose all TelosB nodes to just listen (not talk) to CC2420 radio, there are also other sensor nodes with CC2420 radio which cause 0~35 dB of noise. For channel 11, 0 ~ 3dB of CTI measured. For channel 12, 0~15 dB of CTI measured.

Figure 5.5 shows the RSSI versus distance relation presenting PLE and the RSSI values measured with the same 20 node chain topology in KanseiGenie.

As we can see, channel 26 and channel 12 have the similar average noise floors. However, channel 26 has much higher (about 15 dB) average interference values than channel 12, and accordingly links of channel 26 have much lower (about 18 dB) RSSI values than channel 12. There are frequency varying delays and attenuation for
the individual multi-path components, called frequency selective fading. This causes different RSSI values of the same link for different channels.

Recent studies on WLAN interference on IEEE 802.15.4 (e.g. [18]) validated that interference make very little or no impact on RSSI values. For the long time period, because of the time varying behavior in propagation environment, the overall radio channel is time-variant meaning time-varying delays and attenuations for the individual multi-path components. That causes 2\~3 dB RSSI variances even though noise-floors are stable.

We assume those time-selective and frequency-selective fading are related with interference levels. If the interference level has negative relationship with the link quality, measuring only noise floors enable us to predict the actual link quality of links. Because noise floor measurement can be done in parallel, the measurement time doesn’t scale with the size of the target network. This will ultimately solve the scalability problem of RS-Profiler, we execute and present an experimental study.

Based on the assumption, we initially measure RSSI values of links ($rssi(dB)_{init}$) and the noise floor ($nf(dB)_{init}$). Then, later we only measure the noise floor ($nf(dB)_{reg}$) to predict the RSSI values ($rssi_{reg}$) of links.

Figure 5.5: RSSI vs. distance: KanseiGenie channel 26 (1) and Channel 12 (2)
Therefore, we can predict the rssi value of the specified channel as

\[ rssi(dB)_{reg} = rssi(dB)_{init} + \Delta \]

where \( \Delta \) is \( n_{f_{init}} - n_{f_{reg}} + \beta \), where \( \beta \) is a tuning factor.

### 5.3.2 Validation

We run RS-Profiler to send packets. We measured noise floor (and interference) with 2 KHz frequency, and each packet from the sensor node 16 of the original 4 x 4 grid of KanseiGenie contains 128 noise floor measurement samples. Each packet from node 16 separated about 65 msec, therefore there is about 30 msec time gap between noise floor measurement and the RSSI measurement of the receiving packet. That means we predict the RSSI values of the receiving packet with the noise floor measurement right before (about 30 msec) the reception of it. Figure 5.3 shows the noise floor and the interference measurements in all channels of a node in KanseiGenie. Because the channel 23 is the most interfered one, we execute this study in the channel 23.

![Figure 5.6: RSSI, noise floor and interference measurements of link1 at a node of KanseiGenie in channel 23](image)
Figure 5.7: RSSI, noise floor and interference measurements of link3 at a node of KanseiGenie in channel 23

Figure 5.6 and 5.7 shows the RSSI, noise floor and the interference measurements of link 1 and 3 at a node of KanseiGenie in channel 23 measured about a day long every two hours. RSSI value varies with 3 dB, average noise floor varies within 1 dB, and average and median interference varies within 7 dB in a day. Two links follow the pattern that we expected very weakly, the lower RSSI with the higher interference. But, the variance of RSSI is not as dynamic as the median or average interference is. The average noise floor is also impacted with the interference slightly, negatively related with the RSSI values of links.

Figure 5.8 shows the rssi variances of links measured at a node.

Figure 5.9 shows the cumulative distribution function graphs of errors between the measured RSSI values and the predicted with the static link qualities (1), and with the average noise floor difference (2). Because RSSI values of links do not vary as much as the interference varies, prediction with the assumption of static link quality show very high accuracy (80% or higher accuracy within 2 dB). But, prediction with the average noise floor performs slightly better than that with static
Figure 5.8: RSSI Measurements of links at a node of KanseiGenie channel 23

Figure 5.9: Cumulative RSSI prediction errors at a node of KanseiGenie in channel 23. (1) predicted with static link qualities, (2) predicted with average noise floor difference

link quality assumption. With the prediction of RSSI using noise floor comparison, we can achieve above 90% of accuracy within 2 dB error.

5.4 Accurate prediction of PRR with RSSI values

In this section, we study how to predict PRR, one of the most popular link quality metrics, with RSSI measurement values. If we can accurately predict PRR with RSSI,
then we can reduce $T_{rss}$ by reducing the number of packets to transmit from 100 to much smaller numbers (e.g. 5).

### 5.4.1 Prediction method

It is well known that accurate prediction of PRR in the fast transition area (10% $\sim$ 90%) is very hard, and almost impossible [32], because in the fast transition area 1.5 dB difference can mean 80 $\sim$ 90 % difference in PRR. Nevertheless previous negative results on PRR prediction using SNR, we tried to refine PRR prediction with SNR as best as we could. Liu et. al. [25] argued and validated that SNR-PRR relation equation can vary with location and time. To have more accurate SNR-PRR relation model, we measured the SNR and PRR values in KanseiGenie [2] using RS-Profiler, with 100 ms of inter-packet-interval. Figure 4.1 shows the measured SNR-vs-PRR relations. Figure 4.2 shows the original analytical SNR vs. PRR graph and the experimentally adjusted graph of KanseiGenie. Accordingly we adjusted equation 4.2.1. We use this adjusted SNR-vs-PRR relation to predict the expected PRR of links with measured SNR values in KanseiGenie.

### 5.4.2 Validation

We executed an experimental study with a $4 \times 4$ grid topology in KanseiGenie in channel 11 and with transmission power of -25 dBm. Total 116 links were available for RSSI, noise floor, and PRR measurement.

Figure 5.10 shows the cumulative errors between the measured PRR and the predicted PRR with the measured RSSI, noise floor, and with the adjusted PRR-SNR relation equation. As we can see, there are two different patterns in figure 5.10. With SNR values 8 dB or higher, the PRR prediction is almost perfect (within 1%). But, with SNR values lower than 8 dB, we achieve about 50 % of links within 10
Figure 5.10: Cumulative PRR prediction errors with links of a 4×4 grid network of KanseiGenie in channel 11 with transmission power of -25 dBm.

% error, and about 80 % of links within 20 % error. However, compared to the negative previous studies results (1.5 dB difference can be 90 % difference), our PRR prediction performs very well (PRR of 80 % of links predicted with less than 20 % error).

Figure 5.11: SNR vs. PRR prediction error with links of a 4×4 grid network of KanseiGenie in channel 11 with transmission power of -25 dBm.

Figure 5.11 shows the scatter plot of the relation between SNR values of links and
the PRR prediction error. We can see clearly that for links with 8 dB or higher SNR our adjusted PRR-SNR equation very accurately predict PRR with measured SNR, but for links with SNR lower than 8 dB the PRR prediction performs with 0 ∼ 50 % errors.

5.5 Conclusion

In this chapter, we studied how to relieve the scalability problem of RS-Profiler, that with scale of WSN testbed, profiling time increases over practically allowable limit. The first method is to predict RSSI values of links for different transmission powers based on the measured RSSI values with the highest transmission power. This method is based on the log-normal shading model of the received signal strength. We validated this method with experiments of RS-Profiler over a 4 × 4 grid network of KanseiGenei WSN testbed using 8 different transmission power levels. With the basic prediction method, for more than 90 % of the links, RSSIs are accurately predicted within the error of 2 dB. With the refined prediction method using the RSSI calibration, we can predict RSSI values of about 80% or more links within 1 dB accuracy with just one transmission power level RSSI measurement study.

The second method is to predict the RSSI values of links based on the noise floor measurements by comparing them with the initial noise floor values together with the initial RSSI values. Because noise floor measurements can be done in parallel, this method decreases the profiling time substantially. We validate our method with experiments of RS-Profiler over 9 links of a 4 × 4 grid network of KanseiGenei WSN testbed. With the prediction of RSSI using noise floor comparison, we could achieve above 90 % of accuracy within 2 dB error. However, this method needs further improvement. According to our recent measurement study, in stable channels (e.g. channel 11, 26), noise floor doesn’t vary much (within 0.2 dB) in 3 months, but RSS
values vary up to 5 ~ 10 dB. Therefore, if RSS values had not been updated long
time (e.g. 3 months), this reduction method might work very poorly.

The third method to relieve the scalability problem of RS-Profiler is predicting
the PRR values with the measured SNR values of links. To have more accurate
PRR-SNR relation model, we executed RSSI and noise floor measurement study in
KanseiGenie and collected more than 60,000 PRR-SNR relations, then adjusted the
PRR-SNR relation equation with the collected results. With SNR values 8 dB or
higher, the PRR prediction is almost perfect (within 1%). With SNR values lower
than 8 dB, we achieve about 80 % of links within 20 % error. Compared to the
negative previous studies results (1.5 dB difference can be 90 % difference), our PRR
prediction performs very well.
Achieving performance repeatability of WSN protocols is hard while pursuing the cost and energy efficiency. However, to pursue the well-made and widely adopted WSN applications and services, achieving the repeatable and reliable performance is not elective. In this dissertation, we studied analytical and data driven methods to achieve performance repeatability of WSN protocols working on the cost and energy efficient sensor nodes and systems.

6.1 Achieving repeatable performance of WSN protocols

To achieve the repeatable performance of WSN protocols, we executed and validated the following studies.

- With analytical method, we mathematically characterized the link selection behavior and the expected performance. In chapter 2, we mathematically characterized the link selection behavior (link usage spectrum), and in chapter 3, we characterized the expected performance of WSN routing protocols using 1-hop routing metric (e.g. $PRR \times D$). We validated our mathematical model of link usage spectrum and the expected performance of WSN routing protocols with the vast experimental studies with CTP with ETX and $PRR \times D$ routing
metric on 20 node TelosB chain topologies across indoor corridor, warehouse, and outdoor parking lot.

- We presented two performance matching methods to achieve the consistent performance repetition across environments. In chapter 2, first method matches the link usage spectra of protocols based on our observation that the expected PRR of the links chosen by the protocol is uniformly high and markedly different from the expected PRR of all links at a given distance, especially in the case of long links. In chapter 3, second method matches the expected performance. We validated two methods with extensive experimental studies that executed CTP with ETX and $PRR \times D$ routing metric on 20 node TelosB chain topologies across indoor corridor, warehouse, and outdoor parking lot, and 2-dimensional $4 \times 4$ grid topologies across four major WSN testbeds.

- With data driven method, we simulated wireless protocols with operational models with detailed link quality data collected with a WSN testbed resource specification profiling program, RS-Profiler. In chapter 4, we presented a method to achieve the performance repeatability through an analytical method that can accurately predict the performance of wireless routing protocols, which use 1-hop metric and cumulative metric, based on the measurement of RF resource specification. Because collecting RF data of links may not be feasible in outdoor real deployments, our works will contribute in protocol performance prediction in WSN testbed and performance repetition within and across WSN testbeds.

- In chapter 5, we presented how to relieve the scalability problem of Resource Specification data collection by exploiting that the received RSSI with different transmission powers can be predicted with high probability, and validated with experimental results. We also studied further improvement on data collection
time by reducing the measurement time with accurate prediction on RSSI values based on only noise floor measurement. Our method slightly outperforms the prediction method that assumes link qualities stay constant. We presented a method predicting the PRR values with the measured SNR values of links. With SNR values 8 dB or higher, the PRR prediction is almost perfect (within 1%). With SNR values lower than 8 dB, we achieve about 80% of links within 20% error. Compared to the negative previous studies results (1.5 dB difference can be 90% difference), our PRR prediction performs very well.

- We provided methods to predict the variance and the upper/lower bound of WSN protocol performance. In this dissertation, we presented two methods to predict the performance of WSN protocols. First, we provided a mathematical equation for performance variance which is due to the spatial RSSI variance. Second, we provided accurate performance prediction algorithms using Resource Specification RF data, then applied them to Monte-Carlo simulations. We validated two methods with extensive experiments over 18 4 × 4 grid networks of TelosBs of KanseiGenie. We also provided two methods to deterministically predict the upper/lower bounds of WSN protocol performance, then we validated with experimental results.

6.2 Looking ahead

We are working on publishing our work of achieving performance repeatability across environments with data driven method. The works presented in this dissertation can be extended in several directions. First, the proposed analytic method for predicting link usage provides a good approximation for networks with links of limited temporal variation, but refinements are required to handle networks with links of significant
temporal variation. Second, we can redress the limitations of the link usage spectrum, focusing in particular on providing a continuous version of the concept which allows network nodes to be placed at random points in a geometric space. We can explore predictable performance in the target networks, using knobs other than the transmission power control and taking into account metrics other than those related to the forwarding link selection alone. Third, we can extend our study in this dissertation on the variability of protocol behavior in one network and account for the preservation of that variability in the target network, over and above the average behavior that we have focused on in this work. Fourth, we can study how to extend our technique to adopt more user-specified performance indexes (e.g. goodput, energy efficiency, delay, etc.), and reflect more complex RF environment factors (e.g. macro and micro fading effects). Sixth, we can achieve more accurate RSSI prediction on time, space, and channels. Performance repeatability is a crucial requirement of wireless protocol development, and closely related with debugging process. Resource Specification data collection will provide very precious information in protocol debugging process, and with further refinement and more accurate RSSI prediction will enable very precise and efficient Resource Specification data collection in nearly real-time. However, with only Resource Specification data it will not be enough to root-cause all the bugs and errors, and to reason the wireless protocol behavior thoroughly. Seventh, WSN testbeds equipppped hardware and software solutions that enable us to see the detailed online packet transmission and reception events with the related data (e.g. RSSI, noise floor, collision) and the internal information of sensor nodes (e.g. memory usage, message queue, routing table information) will give us a thorough ability to debug most of problems of WSN protocols, and to reason about the wireless protocol behaviors.
BIBLIOGRAPHY


