Reproducing Consistent Wireless Protocol Performance across Environments

Taewoo Kwon[1], Emre Ertin[2], and Anish Arora[1]

[1] The Ohio State University Columbus, OH 43210, USA
{kwonta,anish}@cse.ohio-state.edu
[2] ertine@ece.osu.edu

Abstract. Full scale experimentation with wireless networks in deployment environments is difficult, so a common validation technique is to test a prototype network in a convenient environment prior to deployment. In this paper, 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 illustrate 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.

Keywords: wireless sensor network, testbed design, wireless protocol performance.

1 Introduction

Experiences in deploying low-power wireless networks during this decade have yielded a number of surprises, wherein network behavior in the field diverged substantially from that seen in laboratory tests. A combination of factors has contributed to these surprises. One key factor is that the effective topology of the laboratory tests is different from that of the field deployment: Not only is inter-node spatial (separation) scale different in the two networks, 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

J. Zheng, D. Simplot-Ryl, V.C.M. Leung (Eds.): ADHOCNETS 2010, LNICST 49, pp. 524–540, 2010.

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. 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 not be 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 [1]. (Power control was realized via attenuator hardware and/or software control.) That study also experimentally compared the performance metrics of two wireless sensor network (WSN) protocols — Sprinkler [2], a protocol that provides a bulk data transmission service, and LOF [3], a protocol that provides a beacon-free routing service— at different spatial scales in Kansei, an indoor WSN testbed [4], [5], 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.

Link Usage Spectrum. In this paper, 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.

The hypothesis of this paper 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.

There are two major factors that affect the link usage spectrum: the metric chosen by the network protocol for selecting parents to who 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 the paper is readily adapted to several other routing metrics.

Related Work. There has been little previous work on the link usage spectrum. There is significant literature however, e.g. [6], 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), PRR × d (the forwarding distance) [7], and expected latency per unit distance ($\mathrm{ELD}=\frac{1}{\mathrm{PRR}\times \mathrm{d}}$) [3]. A related work that implicitly exploits the link usage spectrum idea is [7], 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. [7] also uses numerical simulations for calculating the spectrum; by way of contrast, we provide a closed form equation for expressing the link usage spectrum in the context of the forwarding metric at hand.

Contributions. Our primary contribution includes evidence in support of the hypothesis that link usage spectrum is a gross predictor of network performance. Our evidence consists of experimental results that confirm that the network performance in different settings is most similar when the l_1 distance between the corresponding link usage spectrums is minimum. Specifically, these involve multiple indoor and outdoor experiments with the Collection Tree Protocol (CTP) [8], which is a popular messaging protocol distributed with the TinyOS 2.0 release.

By the same token, i.e., by minimizing the l_1 distance between the two settings, we show a general technique for achieving comparable performance of a network protocol in test and deployment settings.

A third contribution is to show how the link usage spectrum is analytically derived for network protocols. Specifically, we consider the case of protocols whose forwarding metric depends on Packet Reliability Rate (PRR), distance, and other variables based on SNR; this case spans a large fraction of the route selection protocols in use today. Further, we show that the analytically obtained results are corroborated by experimental measurements and simulation of the link usage spectrum in such contexts.

Roadmap. In Section 2, we define link usage spectrum and formulate our method for achieving comparable performance based on the XPlantError metric. In Section 3, we provide an experimental study of CTP protocol transplantation in the context of a simple chain topology in multiple indoor and outdoor environments. In Section 4, we show how to analytically calculate the link usage spectrum for a class of network protocols, and validate the analysis via simulation based results in 1-dimensional and 2-dimensional networks. Finally, we conclude in Section 5, with a discussion of variations of the definition of XPlantError for achieving potentially higher fidelity in predicting network behavior.

2 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:

$$R(d) = P_t - PL(d_0) - 10\eta \log(d/d_0) + N_\sigma \tag{1}$$

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

$$y(d) = R(d) - P_0 \text{ (in dB)} \tag{2}$$

The radio receiver performance can be characterized by representing the packet reception rate $\mathrm{PRR}(y)$ as a function of the received the receiver SNR, y. $\mathrm{PRR}(y)$ gives the probability that a packet received with SNR of y will be decoded correctly by the receiver. The relation between packet-reception-rate and SNR depends on the modulation scheme and the packet encoding scheme employed by the radio transceiver. The function $\mathrm{PRR}(y)$ 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 [9] and analytical [6] 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.

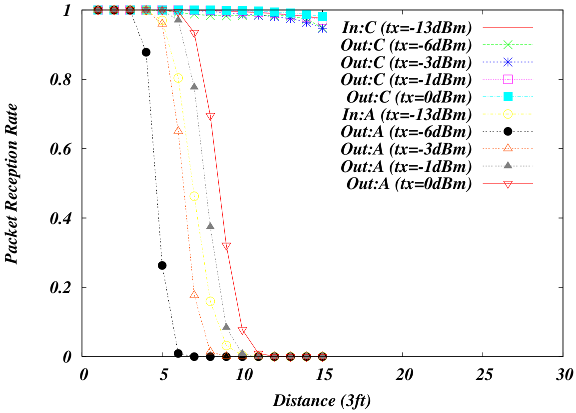

Fig. 1. PRR distribution of links for indoor and outdoor propagation environments at various transmit power levels (C:Chosen, A:All)

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 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 paper, we focus on a particular network statistic called the *link usage spectrum*, which captures *the probability distribution with which a given network protocol selects links of different length from among all available links in the network at hand.*

Example 0. To illustrate the definition of link usage spectrum, consider a wireless network $\mathcal{W} = (\{l_j\}, \eta, \sigma)$ with link set $\{l_j\}_{j=1}^{M}$ and the RF environment (η, σ) 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.

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

$$L(\mathcal{W}, \mathcal{P}, i) = \text{Prob}[d(l) = i\tau] \tag{3}$$

Example 1. We extend the previous example from a one to a two dimensional two dimensional grid network with uniform node spacing. In this case, the link lengths $d_j \equiv d(l_j)$ and the progress to destination $p_j \equiv p(l_j)$ are constrained to the finite set $\{(\tau,\tau), (\tau, 2\tau), (\tau, 3\tau), \ldots, (\tau, N\tau),\ (2\tau,\tau), (2\tau, 2\tau), (2\tau, 3\tau), \ldots, (2\tau, N\tau), \ldots, (N\tau,\tau), (N\tau, 2\tau), (N\tau, 3\tau), \ldots, (N\tau, N\tau)\}$, where τ is the minimum node spacing. We index elements of the above finite set from 1 to M. The link usage spectrum $L(\mathcal{W}, \mathcal{P}, i)$, then, is the discrete probability distribution over the length of links and the progress of links to the destination induced by the network protocol $\mathcal{P}$.

$$L(\mathcal{W}, \mathcal{P}, i) = \text{Prob}[d(l) = d_i, p(l) = p_i], \tag{4}$$

where i is the index of $(m\tau, n\tau)$, $d_i = \sqrt{(m\tau)^2 + (n\tau)^2}$, $p_i = d_i * cos(|tan^{-1}(\frac{n}{m}) - \theta|)$, θ 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 5. As a result, we propose to use the link usage spectrum to match protocol behavior across scales and environments.

2.1 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 spectrums of the two networks; one way of realizing this is by optimizing the selection of the transmit power.

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

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

XPlantError is essentially the l_1 distance between the link usage spectrums for the two networks $\mathcal{W}$ and $\tilde{\mathcal{W}}$ which differ in scale and RF environment. Theorems 1 and 2 show the relation between the link usage spectrum and 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 transmit power in the network $\tilde{\mathcal{W}}$ is variable through $\tilde{P_t} = \tilde{P_0} + \beta$ where β is the power attenuation or amplification in the scaled network. Since the scaled vector in general reduces the node distances for convenient testing, β in general is a negative value indicating power attenuation. As a result the scaled network realizations $\tilde{\mathcal{W}}(\beta)$ depends on β[1]. The optimal power attenuation is then chosen to minimize the XPlantError metric:

$$\beta_{\text{opt}} = \arg\min_{\beta} \text{XPlantError}(\mathcal{W}, \tilde{\mathcal{W}}(\beta), \mathcal{P}) \tag{5}$$

As shown in [1], if the two networks are in the same environment (i.e., $\eta = \tilde{\eta}$ and $\sigma = \tilde{\sigma}$)) then β 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 β in that case would result in identical link usage spectrums. 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.

3 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) [8]. 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 transmit powers are chosen to minimize the distance between link usage spectrums for test and deployment environments.

3.1 Experimental Setup

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

---

[1] We could also introduce spatial variations in β across the nodes to influence the width of the resulting spectrum, for a better match.

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 β of (7dB, 10dB, 13dB); and three output levels (-18 dBm, -13 dBm and -10 dBm) for an indoor network in an open warehouse.

3.2 Physical and Link Layer

The CC2420 [11] 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:

$$\mathrm{BER} = Q(\sqrt{2y/\mathrm{PG}/\mathrm{CG}}) = \frac{1}{2}\mathrm{erfc}(\sqrt{y/\mathrm{PG}/\mathrm{CG}}), \tag{6}$$

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 [12]. Thus, the Packet Reception Rate equation is calculated to be

$$PRR = \left(1 - \frac{1}{2}\mathrm{erfc}(\sqrt{(x - 11 - P_0\ \mathrm{dB})})\right)^{8*packet\_size} \tag{7}$$

where x denotes the SNR in dBm. We note that the bit-error-rate approximation given in Equation 6 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 $\sim$ 20 unit (1 unit= 3ft)) within maximum communication range with the highest transmission power level (0dBm). For the outdoor test, 10 measurements were taken from the distance of 1 $\sim$ 10 unit distances where 30 ft seems to be the maximum communication range with the same transmission power, 0 dBm. Figure 2 shows the observed received signal values and the associated log-normal fit. Table 1 presents the summary.

3.3 Messaging Layer

CTP [9] is a tree-based collection protocol. Nodes generate routes to the sink using a routing gradient. CTP uses ETX [13] as the default routing metric.

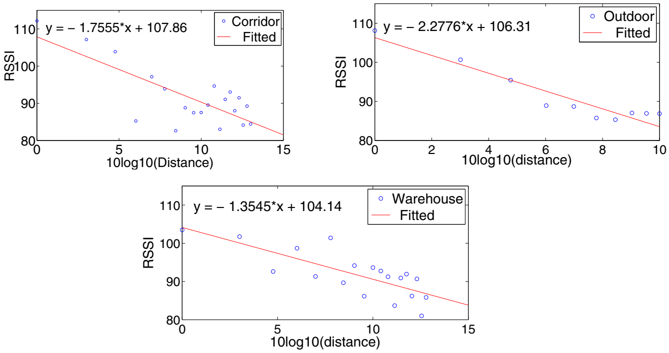

Fig. 2. Indoor Corridor, Outdoor & Indoor Warehouse (RSSI vs. Distance)

Table 1. Log normal model variables for Indoor Corridor, Outdoor, and Indoor Warehouse RF environments

Metrics	Indoor Corridor	Outdoor	Indoor Warehouse
Path Loss Exponent	1.7555	2.2776	1.3545
RSSI Standard Dev.	5.0 dB	4.2 dB	5.0 dB

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 [7].

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 trough 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.

3.4 Results

We first present link usage spectrum for indoor and outdoor environments at various transmit power levels using data from the experiments in Figure 3. We see that the XPlantError is minimized for the corridor and the outdoor experiments when the latter uses a transmit power level between -3dBm and 0dBm, and for the corridor and warehouse experiments when the latter uses a transmit power level between -18dBm and -13dBm. Table 2 shows that the outdoor XPlantError is minimized for the 0dBm case, and the warehouse XPlantError is minimized for the -18dBm case.

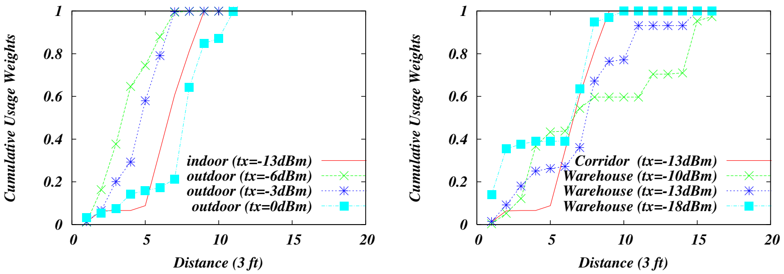

Fig. 3. Cumulative Link Usage Spectrum (Left: Corridor vs Outdoor, Right: Corridor vs Warehouse)

Table 2. Comparison of l_1 distance between Corridor network with tx=-13dBm and various Outdoor/Warehouse networks

l_1 distance	Outdoor tx=-6dBm	Outdoor tx=-3dmB	Outdoor tx=0dBm	Warehouse tx=-10dBm	Warehouse tx=-13dBm	Warehouse tx=-18dBm
Anlytical	2.7556	1.8767	1.203	3.2463	1.6976	1.5986

Next, we show that matching link usage spectrums results in consistent protocol performance as measured by commonly adopted metrics of mean-link-length and end-to-end latency in these experiments. Table 3 and 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 transmit power level between -3dBm and 0dBm, and by the warehouse environment network when the latter uses a transmit power level between -18dBm and -13dBm, which is consistent with our main hypothesis.

Table 3. Link Length Statistics for Corridor, Outdoor, and Warehouse Experiments

Link Length	Indoor tx=-13dBm	Outdoor tx=-6dBm	Outdoor tx=-3dBm	Outdoor tx=0dBm	Warehouse tx=-10dBm	Warehouse tx=-13dBm	Warehouse tx=-18dBm
Average	6.9329	4.1842	5.0584	7.7749	8.6506	7.6289	5.4536
Median	7	4	5	8	8	7	5

Table 4. End-to-End Delay for Corridor, Outdoor, and Warehouse Experiments

e-to-e delay	Indoor tx=-13dBm	Outdoor tx=-6dBm	Outdoor tx=-3dBm	Outdoor tx=0dBm	Warehouse tx=-10dBm	Warehouse tx=-13dBm	Warehouse tx=-18dBm
Average	3.712	5.2481	4.4512	3.2049	2.2291	2.6072	4.2544
Median	4	5	4	3	2	3	4

4 Analytic Methods for Predicting Link Usage Spectrum

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 PRR×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 theorems here are presented using PRR×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.

4.1 Analytic Model for 1-Dimensional and 2-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 multidimensional 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 × d as the metric for choosing forwarding links, the probability of choosing link l_i over link l_j is expressed as follows:*

$$
\begin{aligned}
&P(PRR(y_i) * p_i > PRR(y_j) * p_j) \\
&= \int_{-\infty}^{a} \frac{1}{\sigma\sqrt{2\pi}} \left(\frac{1}{2} + \frac{1}{\sqrt{\pi}} \left(z - \frac{z^3}{3} + ...\right)\right) e^{-\frac{(y_i-\mu_{y_i})^2}{2\sigma^2}} dy_i \\
&+ \int_{a}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \left(\frac{1}{2} + \frac{1}{2} erf\left(\frac{g(\beta_j) - \mu_{y_j}}{\sqrt{2}\sigma}\right)\right) e^{-\frac{(y_i-\mu_{y_i})^2}{2\sigma^2}} dy_i,
\end{aligned}
$$

where $z = \frac{y_i - \mu_{y_j}}{\sigma\sqrt{2}}$*, and* y_i, y_j *are SNR (dB) values of link* l_i*,* l_j*, and* p_i, p_j *are progresses to destination for link* l_i*,* l_j*,* $\beta_j = \frac{p_j}{p_i}$*,* $a = PRR^{-1}(\min\{\beta_j, \frac{1}{\beta_j}\})$*,* $g(\beta_j) = a$ *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 $PRR \times D$,

$$\text{Prob}(PRR(y_i) * d_i > PRR(y_j) * d_j)$$

$$= \text{Prob}\left(PRR(y_i) > \beta_j PRR(y_j)\right)$$

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:

$$PRR(y_i) > \beta_j PRR(y_j) \approxeq y_i > h(y_j), \tag{8}$$

where

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

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

$$\begin{aligned} &\text{Prob}\left(PRR(y_i) > \beta_j PRR(y_j)\right) \\ &= \text{Prob}(y_i > y_j, y_j < a) + \text{Prob}(y_i \geq a, y_j \geq a) \\ &= \text{Prob}(y_i > y_j, y_i < a) + \text{Prob}(y_i \geq a) \end{aligned}$$

Because y_i, y_j are Gaussian:

$$\begin{aligned} &\text{Prob}\left(PRR(y_i) > \beta_j PRR(y_j)\right) \\ &= \int_{-\infty}^{a} \frac{1}{\sigma\sqrt{2\pi}} \left(\frac{1}{2} + \frac{1}{2} erf\left(\frac{y_i - \mu_{y_j}}{\sigma\sqrt{2}}\right)\right) e^{-\frac{(y_i - \mu_{y_i})^2}{2\sigma^2}} dy_i \\ &+ \int_{a}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \left(\frac{1}{2} + \frac{1}{2} erf\left(\frac{g(\beta_j) - \mu_{y_j}}{\sqrt{2}\sigma}\right)\right) e^{-\frac{(y_i - \mu_{y_i})^2}{2\sigma^2}} dy_i, \end{aligned}$$

where

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

Finally, we arrive at the expression in Theorem 1 retaining the first two terms of the Taylor series of $erf(z)$, $erf(z) = \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n+1}}{n!(2n+1)}$.

Theorem 2. *For a protocol* $\mathcal{P}$ *which uses* $PRR \times d$ *as the metric for choosing forwarding links, the probability of choosing link* l *of length over all other links is expressed as follows:*

$$L(\mathcal{W},\mathcal{P},index(l)) = P\left[PRR(y(l))*p(l) = \max_{j=1\ldots M}\{PRR(y_j)*p_j\}\right]$$

$$= \sum_{k=0}^{M-1}\int_{a_k}^{a_{k+1}}\frac{1}{\sigma\sqrt{2\pi}}\left[\prod_{j=1}^{k}\left(\frac{1}{2}+\frac{1}{2}erf\left(\frac{g(\beta_j)-\mu_{y_j}}{\sqrt{2}\sigma}\right)\right)\right]$$

$$\left[\prod_{j=k+1}^{M}\left(\frac{1}{2}+\frac{1}{\sqrt{\pi}}\left(z_j-\frac{z_j^3}{3}+\ldots\right)\right)\right]e^{-\frac{(y-\mu_{y(l)})^2}{2\sigma^2}}dy,$$

where $\beta_j = p_j/p(l)$, $a_j = PRR^{-1}(\min\{\beta_j, \frac{1}{\beta_j}\})$, *and the links are enumerated such that* $a_1 \le a_2 \le \ldots \le a_{M-1}$, *with* $a_0 = -\infty, a_M = \infty$ *and* y_j *is the SNR experienced by link* l_j, $z_j = \frac{y_j-\mu_{y_j}}{\sigma\sqrt{2}}$, *and*

$$g(\beta_j) = \begin{cases} a_j, & \text{if } \beta_j \ge 1 \\ \infty, & \text{if } \beta_j < 1 \end{cases}$$

The proof is an immediate extension of Theorem 1 and is given in [14]. Both results rely on the approximation used in Equation 8 to convert the nonlinear boundary where one link is preferred over another to a piece-wise linear boundary suitable for close-form integration.

4.2 Validation of Analytical Approximations

We performed Monte-Carlo simulations to verify the accuracy of this analytical approximation of the link usage spectrum for various radio propagation environments and transmit power levels. 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 in the experiments described in the previous section. We observed that although the general behavior of link usage spectrum as a function

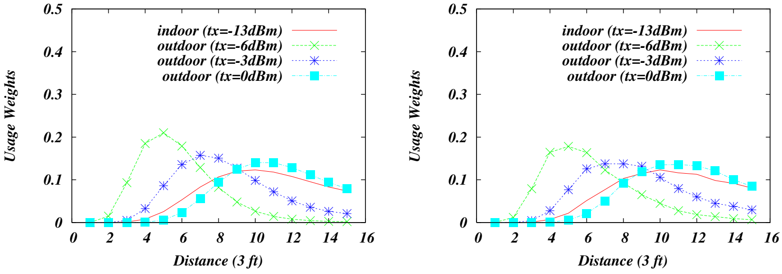

Fig. 4. Link Usage Spectrums. Left: Analytical (Theorem 2), Right: Monte-Carlo Simulation

of transmit 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 suggested that temporal RSSI variations are one major 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 1, if we only consider spatial RSSI variations, PRR of all the chosen links are close to 1. However, if we consider temporal RSSI variations (typically, with a standard deviation of 3dB), long links will suffer in greater degree because they are likely to be closer to the threshold SNR and therefore subject to major fluctuations in their PRR values. As a result, the routing protocol will deselect long links encountering temporal fades. To account for this observed behavior, we consider only links whose unconditional (prior) expected PRR is larger than 1%. For example, using figure1, for an outdoor network with $0dB$ transmission power, this would correspond to all links whose length is less than or equal to 10. The result is shown in Figure 5.

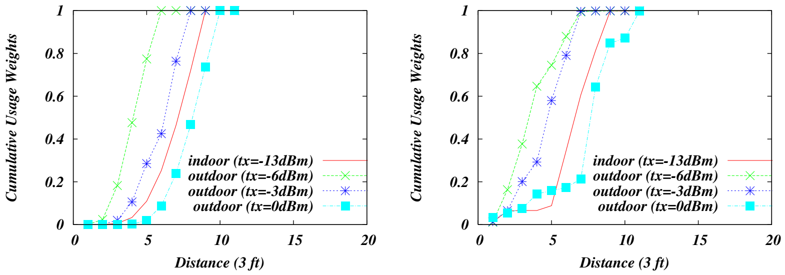

Fig. 5. Cumulative Link Usage Spectrum (Left: Theorem 2 with Refinement, Right: Experiments)

2-Dimensional Simulation Study. We also compared analytical and simulation results of the link usage spectrum for the CTP protocol in a two dimensional 10×10 grid topology. The internode distance in this case is 6ft, and the sender is located at the leftmost bottom corner, and the destination is at the rightmost top corner. We use Theorem 2 to calculate the two dimensional link usage spectrum, and use TOSSIM 2 [15] simulator to obtain empirical data to validate analytical results. The source node is located at grid location (1,1) and it sends packets over the multihop network to the sink node located at (10,10).

Figure 6 shows the analytical results calculated with Theorem 2. Figure 7 shows TOSSIM 2 [15] simulation results. We note that analytical expressions assume an infinite network model, calculate performance per unit distance, and scale the results to a finite size network. Analytical results show that a given indoor network of transmission power (-17dBm) is matched optimally with outdoor network of (-5dBm) transmission power, by minimizing the XPlantError.

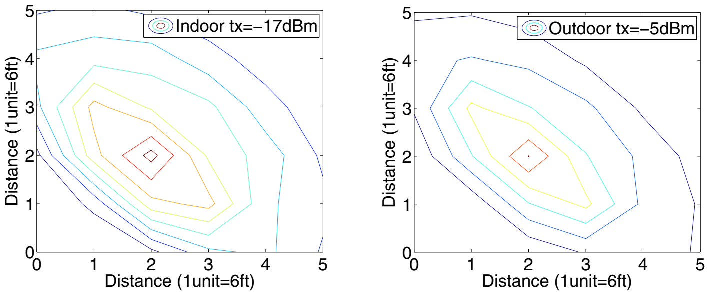

Fig. 6. Link Usage Spectrum (Analytical). Left: Indoor -17dBm and Right: Outdoor -5dBm

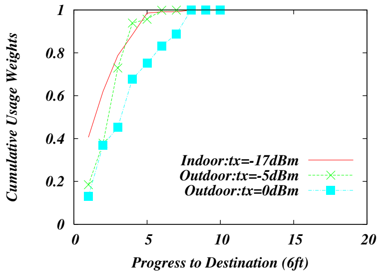

Fig. 7. Cumulative 2-Dimensional Link Usage Spectrum (TOSSIM 2 Simulation)

Figure 6 shows contour plots for 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 $-12dB$.

The simulation results of these two networks closely follow the analytical results. In Figure 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 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 5 shows average end-to-end delays and link progresses for each environments. As predicted with analytical results, indoor -17dBm matches outdoor -5dBm better than it does 0dBm.

Table 5. End-to-end Delay and Link Progress for 2-D Environments (TOSSIM 2 Simulation)

	Indoor tx-17dBm	Outdoor tx=0dBm	Outdoor tx=-5dBm
e-2-e delay	5.23	3.85	4.68
link progress	2.34	3.9	2.81

5 Conclusion and Future Work

In this paper, we studied a method for achieving comparable protocol performance across deployment and test environments. We have found the method to be valid across several protocols (which in turn were based on different forwarding metrics), various performance metrics, and diverse environments and presented some of those results here. We have also shown that our analytic methods for predicting link usage yielded good approximation in some environments, but need refinement in others (for instance, when temporal variation of links was significant).

Error metrics other than XPlantError can be employed for matching performance. In particular, one can choose the transmit 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:

$$\left| \sum_{i=1}^{n} g(y_i | l_i \uparrow) L(W, \mathcal{P}, i) - \sum_{i=1}^{n} g_{(}\tilde{y}_i | l_i \uparrow) L(\tilde{W}(\beta), \mathcal{P}, i) \right| \tag{9}$$

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 spectrums proposed in this paper. Also, in this paper, we gave evidence that the use of the generic method of link usage spectrum matching suffices to obtain comparable performance over a wide variety of network metrics. Figure 8 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.

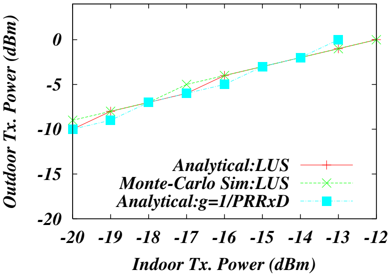

Fig. 8. Relation between Indoor Corridor and Outdoor transmit power for minimizing l_1 distance between the link usage spectrums. Also, relation between Indoor Corridor and Outdoor transmit power to match end-to-end latency performance (cyan).

References

1. Naik, V., Ertin, F., Zhang, H., Arora, A.: Wireless testbed bonsai. In: 2nd International Workshop on Wireless Network Measurement (2006)
2. Naik, V., Arora, A., Sinha, P., Zhang, H.: Sprinkler: A reliable and energy efficient data dissemination service for wireless embedded devices. In: 26th IEEE Real-Time Systems Symposium (2005)
3. Zhang, H., Arora, A., Sinha, P.: Learn on the fly: Data-driven link estimation and routing in sensor network backbones. In: 25th IEEE International Conference on Computer Communications (2006)
4. Arora, A., Ertin, E., Ramnath, R., Nesterenko, M., Leal, W.: Kansei: A high-fidelity sensing testbed. IEEE Internet Computing 10(2), 35–47 (2006)
5. Kansei Testbed, `http://exscal.nullcode.org/kansei`
6. Zuniga, M., Krishnamachari, B.: Analyzing the transitional region in low power wireless links. In: IEEE SECON (2004)
7. Seada, K., Zuniga, M., Helmy, A., Krishnamachri, B.: Energy-efficient forwarding strategies for geographic routing in lossy wireless sensor networks. In: Sensys (2004)
8. TEP 123, `http://www.tinyos.net/tinyos-2.x/doc/html/tep123.html`
9. Zhao, J., Govindan, R.: Understanding packet delivery performance in dense wireless sensor networks. In: Sensys (2003)
10. C. Technology, "Telosb datasheet", `http://www.xbow.com/Products/Product_pdf_files/Wireless_pdf/TelosB_Datasheet.pdf`
11. Chipcon, "cc2420 datasheet", `http://www.chipcon.com/files/CC2420_Data_Sheet_1_4.pdf`
12. Lanzisera, S., Pister, K.: Theoretical and practical limits to sensitivity in ieee 802.15.4 receivers. In: IEEE International Conference on Electronics, Circuits and Systems (2007)
13. Couto, D., Aguayo, D., Bicket, J., Morris, R.: A high-throughput path metric for multi-hop wireless routing. In: Mobicom (2003)
14. Kwon, T., Ertin, E., Arora, A.: Transplanting protocols across different environments. Tech. Rep. OSU-CISRC-9/08-TR47, Ohio State University (2008)
15. TOSSIM 2, `http://docs.tinyos.net/index.php/TOSSIM`