In Defense of Wireless Carrier Sense

by

Micah Z. Brodsky

B.S. Computer Engineering

University of Washington, 2005

Submitted to the Department of Electrical Engineering and Computer Science in partial fulfillment of the requirements for the degree of Master of Science in Electrical Engineering and Computer Science

at the

Massachusetts Institute of Technology

February 2009

Signature of author: _____

Department of Electrical Engineering and Computer Science
December 11, 2008

Certified by: _____

Robert T. Morris
Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by: _____

Terry P. Orlando
Professor of Electrical Engineering and Computer Science
Chair, Department Committee on Graduate Students

In Defense of Wireless Carrier Sense

by

Micah Z. Brodsky

Submitted to the Department of Electrical Engineering and Computer Science on December 11, 2008 in partial fulfillment of the requirements for the degree of Master of Science in Electrical Engineering and Computer Science at the Massachusetts Institute of Technology

Abstract

Carrier sense, or clear channel assessment (CCA), is widely used in wireless medium access control (MAC) protocols as the means to arbitrate access and regulate concurrency, striking a balance between interference protection and spatial reuse. Criticized widely in the literature, carrier sense has been subject to many replacement attempts with sophisticated and complex alternatives, yet it remains extremely popular. Is the search for a superior alternative justified? In this thesis, we develop a physically motivated theoretical model for *average case* carrier sense behavior in the two-sender case, based upon radio propagation theory and Shannon capacity. We argue from our model that common notions about carrier sense, such as the hidden and exposed terminal phenomena, are inherently misleading in the context of adaptive bitrate, casting in black and white terms effects that often cause only mild reduction in throughput. The frequency of severe misbehavior is low. We also demonstrate that it is possible to choose a fixed sense threshold which performs well across a wide range of scenarios, in large part due to the role of the noise floor. The noise floor has a significant effect on fairness as well. Using our model, we show that, when implemented well, average-case carrier sense performance is surprisingly close to optimal. We conclude with experimental results from our indoor 802.11 testbed, which corroborate these claims.

Thesis supervisor: Robert T. Morris
Title: Professor of Electrical Engineering and Computer Science

1. Introduction

Carrier sense is at the heart of many a successful wireless MAC protocol, yet it has long been known to be an imperfect solution. Carrier sense uses measurements of the wireless channel at the senders to decide whether to transmit, but at the physical layer, it is the channel conditions at the receivers that are ultimately relevant. When all nodes are tightly clustered in a compact, isolated network, this might be expected to work very well, because channel conditions at all the nodes are highly correlated. But, for any network broad enough to contemplate spatial reuse, the use of carrier sense has been suspect. Two classic failures have been described, the exposed terminal problem, where nodes that could have transmitted concurrently do not, sacrificing potential concurrency, and the hidden terminal problem, where nodes that shouldn't transmit concurrently do, potentially destroying one or both transmissions.

Concern over carrier sense's weaknesses has inspired a long line of research (e.g. [Bhargavan94], [Jamieson05], [Vutukuru08], [Gollakota08]), and indeed, the investigations in this thesis grew out of our attempts to contribute to this area. We were particularly interested in modern, multi-rate radios, as prior work almost invariably restricted itself to a single bitrate, even though bitrate adaptation is at present the single most important factor in performance under the MAC's control. But our attempts at exploiting hidden and exposed terminals proved consistently disappointing. There were certainly cases where carrier sense made the wrong choice, but nothing seemed significant enough to offer meaningful room for improvement in overall throughput. Could it be that, in practice, carrier sense was already performing close to optimal?

From a worst-case standpoint, it is clear that carrier sense *can* misbehave. But, little research addresses the question, "Is it bad enough to matter?" In particular, how *often* does carrier sense make a poor decision, and how much throughput is sacrificed as a result?

In this thesis, we quantitatively analyze the throughput efficiency of carrier sense. In pursuit of generality, our main results come from analytical modeling, based on standard statistical radio propagation models from the EE community, reasonable assumptions about network layout, and Shannon's capacity formula, which we use as a rough approximation of practically achievable throughput for an adaptive bitrate radio. We show that the answer to the questions posed above, the amount of throughput carrier sense loses, is surprisingly small.

The key intuition behind our results is that interference is a global phenomenon, affecting to some extent all nodes everywhere. The fact that a sender is communicating with a given receiver in the first place automatically puts some bounds on how different channel conditions are at the two nodes; carrier sense can only be so far off the mark. With adaptive bitrate, a node can adjust to whatever particular level of noise and interference it experiences, including conditions slightly less than optimal. Catastrophic mistakes, where interference from concurrent transmissions forces bitrates close to zero, are rare. Conversely, when senders unnecessarily multiplex transmissions, they usually don't lose much throughput, because when each one's turn to transmit arrives, the channel will be clearer and higher bitrates will be feasible. Carrier sense's decisions are not perfect, but given the flexibility of adaptive bitrate, they are quite reasonable. Further tweaking offers only limited benefits – our theoretical analysis indicates average throughput is typically less than 15% below optimal. We also present experimental results from our 802.11 testbed, which reinforce the same conclusions.

This work presents three main contributions. First, a demonstration that carrier sense, while imperfect in theory, provides nearly optimal throughput in the common case, and an analysis identifying the underlying causes for this good behavior. Second, the identification of several distinct behavioral regimes for carrier sense, most of which perform well and only one of which typically encounters bad behavior. Finally, a simple but useful model for the high-level properties of adaptive bitrate radio throughput, focused not an abstracted worst case at the MAC layer but on average-case behavior under realistic radio propagation, whose conclusions in this analysis are confirmed by experiment.

2. Modeling radio capacity

The effectiveness of carrier sense ultimately hinges on the properties of radio propagation. If radio were like a wire, delivering a strong signal everywhere, there would be no question – carrier sense would be ideal. And if radio were abrupt and unpredictable, delivering a signal for a certain distance and then sharply falling off, or casting dark and concealing shadows around obstacles, carrier sense could perform quite poorly. The truth, of course, is somewhere in between. Fortunately, there is no need to speculate; a wealth of research is available quantifying and modeling the rates of signal decay and its variation from place to place (e.g. [Vaughan03], [COST231]).

For modeling radio propagation, we employ the basic path loss - shadowing - fading model (see the appendix for a short overview; [Akaiwa97] also gives a brief but accessible introduction, and a deeper treatment can be found in, e.g., [Vaughan03]). In brief, the model breaks radio propagation into three components: "path loss", a deterministic power law decay with distance, "shadowing", a lognormal statistical variation from place to place due to obstacles and reflections, and "fading" (also called "fast fading" or "multipath fading"), a fine-grained Rayleigh or Rician statistical variation in both space and frequency.[1] The path loss power law exponent α typically varies from 2 to 4 [Vaughan03 p166] [ITU-R P.1238], and shadowing standard deviation σ is typically in the realm of 4-12dB [Akaiwa97] [ITU-R P.1238] [COST231 §4.7.6].[2] We restrict our attention mainly to wideband channels, employed by most modern packet radio specifications including 802.11, which allows us largely to average fading away.

To apply this model to the MAC layer, we need a means to estimate average achievable throughput given signal power, interference power, and a reasonable bitrate adaptation algorithm (such as [Bicket05]). For this purpose, we employ the Shannon capacity formula Capacity / Bandwidth(Hz) = log(1 + SNR), which represents a theoretical upper bound but in practice can be used as a rough proportional estimate.[3]

[1] For indoor networks in commercial buildings with heavy, uninterrupted floors (e.g. slab reinforced concrete), it is wise to introduce an additional term, separate from shadowing, to model floor attenuation [ITU-R P.1238]. We do not explicitly consider multi-floor networks in this thesis.

[2] Applied to our own indoor 802.11 testbed at 2.4GHz, we find $\alpha \approx 3.5$, $\sigma \approx 10$dB and a reasonably good fit. See Figure 14 in the appendix.

[3] We assume interference and background noise can be treated the same, which is theoretically reasonable (though not exact) and borne out by our experiments.

3. Theoretical model of carrier sense

3.1. Introduction

At a high level, MACs rely on carrier sense to answer one fundamental question: "Can I transmit now?" In more detail, the problem breaks down into two questions: "Is my intended receiver listening?" and "Is it a good idea to transmit now, based on interference?" In this thesis, we focus on the latter.

There are several different ways of implementing "carrier sense" in radio hardware. The simplest approach is to continuously monitor power level on the wireless channel and to postpone all transmission attempts while the power is above some threshold. This is combined with a set of MAC-level scheduling and back-off mechanisms in order to approximate fair multiplexing of transmission attempts among mutually-sensing contenders for the channel. In practice, several drawbacks and corner cases associated with this method lead to common standards (e.g. 802.11 [IEEE07]) and common hardware relying on somewhat more complex implementations. Typical elaborations include scanning specifically for packet preamble signatures and measuring the power therein [IEEE07] [Aoki06] [Heiskala02], active measurement of the noise floor as a reference point [Heiskala02], and sensitivity to fields in the demodulated MAC header. The common thread, though, is comparison of a power metric against some fixed threshold, independent of the identities of the sender and receiver and independent of the details of the packet.

For the purposes of this thesis, we model carrier sense as making a simple binary choice between concurrency and multiplexing, abstracting away the details of scheduling. For the sake of tractability, we only consider the common case of two competing sender-receiver pairs. Both senders simultaneously agree to transmit concurrently or to take turns, depending on the (equal) signal power levels each senses from the other compared to some fixed threshold.

When two independent senders compete for the medium, carrier sense lacks the information necessary to make a theoretically optimal decision. This would require four signal powers (S1→R1, S1→R2, S2→R1, S2→R2) plus noise floor information, none of which are conveniently available. Instead, carrier sense uses the conveniently available S2→S1 power as a proxy, deferring when power is above some threshold and transmitting concurrently when below. Thus, the questions we must resolve are, how good is this approximation in practice, and where does the threshold come from? We will investigate these questions using our theoretical model.

3.2. Formal model

3.2.1. Overview

In order to address the above questions, we construct an abstract model of MAC throughput under different concurrency policies, comparing the performance of carrier sense against pure concurrency, time-division multiplexing, and "optimal". As we are interested in "typical" or "average" behavior, we will consider expected throughput averaged over all configurations of a network. The basic scenario is that in Figure 1: two sender-receiver pairs competing for the medium. Assigning precise position coordinates gives a particular "configuration". In our model, nodes communicate with a neighbor found at some random location within a given radius in two-

dimensional space. We refer to this maximum radius, R_{max}, as "network range", not because it is some maximum range of transmission or interference but because it is the maximum range of interest, beyond which either capacity is too low to be useful or no nodes of interest are clustered. Next, given two such sender-receiver pairs, with the senders separated by some specified distance D, we can calculate expected total throughput by averaging over all configurations consistent with free parameters D and R_{max}. Network range R_{max} and sender-sender distance D are the two network-defining parameters in our model, for which we do not assume any prior distribution; it will turn out that they are key in determining network efficiency and carrier sense behavior.

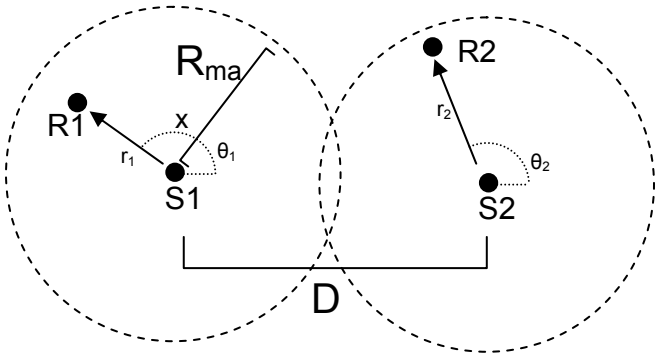

Figure 1 – Model scenario.

We make several assumptions:

- We only consider n = 2 simultaneously contending senders. Small n > 2 does not appear to fundamentally alter the results, but it does complicate matters dramatically, and measurement studies such as [Cheng06] argue that it is of limited importance.

- We assume omnidirectional antennas at all nodes.

- We assume nodes are able to achieve capacity following the rough shape of Shannon capacity (less by some constant fraction) through bitrate adaptation.

- We assume that receivers are distributed independently and approximately uniformly within R_{max} of their senders.[4]

For situations where nodes have a variety of possible receivers scattered around them (e.g. busy access point networks and mesh networks), our average directly models average sender throughput. In other situations, such as residential access point networks, our average is simply an ensemble average, predicting an average over the ensemble of layouts seen in practice.

3.2.2. Analytical formulation

To formulate our capacity model analytically, we place the first sender at the origin and allow its receiver to be at any polar coordinates (r, θ) with $r < R_{max}$. We then place the second sender, the "interferer", on the -x-axis at (D, π).

[4] This does exclude multi-hop paths (3 or more hops) driven close to saturation, because receivers are more likely to lie close to the line segment connecting concurrent senders.

We model carrier sensing capacity as a piecewise function, switching between concurrent transmission and time-division multiplexing ("multiplexing") depending on the interferer signal strength. Thus, we will develop expressions for concurrent and multiplexing capacity first.

For the cases of concurrent transmission, multiplexing, and carrier sensing, the location of the interferer's intended receiver does not matter to our sender – the allocation of channel resources and distribution of interference are unchanged. These cases are symmetric, allowing us to consider throughput one sender-receiver pair at a time.

On the other hand, any notion of "optimal" must consider both sender-receiver pairs because, for example, there may be a need for trade off a small loss at one for a large gain at the other. Our notion of optimal will be maximum aggregate throughput across both links, subject to a weak fairness constraint: both senders must receive equal channel resources, either in the form of equal, exclusive transmit time or equal power under concurrency. This prevents "optimal" from completely starving the weaker link in order to serve the stronger one exclusively, preserving a fairness notion consistent with common carrier sensing MAC policies. In other words, our hypothetical optimal MAC must choose whether to transmit concurrently or to multiplex equally between the two senders, whichever maximizes the resulting sum of throughputs. This is a binary decision which inherently depends on the positions of both receivers, as well as the particular shadowing and fading environments they find themselves in.

The following formulas give an individual sender-receiver pair's throughput, which we will subsequently average over $r = 0..R_{max}$, $\theta = 0..2\pi$.

Using Shannon capacity as our capacity model, a single sender without any competition has a throughput proportional to:

$$C_{single}(r, \theta) = \log(1 + P_0 \cdot r^{-\alpha} \cdot L_\sigma / N_0)$$

Where P_0 is the signal power at unit distance, L_σ is a random variable drawn from the lognormal shadowing distribution, and N_0 represents the thermal noise floor in the channel. Without loss of generality, we can factor P_0 into the noise term. This gives:

$$C_{single}(r, \theta) = \log(1 + r^{-\alpha} \cdot L_\sigma / N)$$

For most calculations, we will use -65dB for $N = N_0 / P_0$.[5] The precise value of N is not important; changing the power level (or noise floor) is equivalent to rescaling the distances. For this power level, r = 20 gives roughly 26dBm SNR, which is reasonable for 802.11a/g 54Mbps. We will run most of our analyses out to r = 120, corresponding to an SNR just shy of 3dB, which is about the minimum practical for any useful connectivity with 802.11 1Mbps.

When two senders of equal transmit power are present, an ideal TDMA MAC gives each a throughput of

$$C_{multiplexing}(r, \theta) = \log(1 + r^{-\alpha} \cdot L_\sigma / N) / 2$$

[5] -65dB is convenient for networks like 802.11 (which transmit around 15dBm and have a noise floor around -95dBm), because empirically, it scales r = 1 to be roughly a human-scale distance from the transmitting antenna.

If instead, the senders transmit concurrently, they each get throughput according to

$$C_{concurrent}(r, \theta) = \log(1 + r^{-\alpha} \cdot L_\sigma / [N + L'_\sigma \cdot (\Delta r)^{-\alpha}])$$

where $\Delta r = \sqrt{[(r \cdot \cos(\theta) + D)^2 + (r \cdot \sin(\theta))^2]}$ (distance between interferer and receiver) and where L_σ and L'_σ are independent values drawn from the same lognormal shadowing distribution.

For a carrier sensing MAC,

$$C_{cs}(r, \theta) = \begin{cases} C_{multiplexing}(r, \theta) & D^{-\alpha} \cdot L''_\sigma > P_{threshold} \\ C_{concurrent}(r, \theta) & D^{-\alpha} \cdot L''_\sigma < P_{threshold} \end{cases}$$

for some specified threshold power $P_{threshold}$. Often, we will find it instead convenient to speak in terms of a threshold distance, $D_{threshold} = P_{threshold}^{1/\alpha}$ which, in the absence of shadowing, is the separation distance at which the two senders begin to transmit concurrently.

An optimal MAC would achieve an average throughput of:

$$C_{max}(r_1, \theta_1, r_2, \theta_2) = \tfrac{1}{2} \cdot \mathrm{Max}[C_{concurrent}(r_1, \theta_1) + C_{concurrent}(r_2, \theta_2), C_{multiplexing}(r_1, \theta_1) + C_{multiplexing}(r_2, \theta_2)]$$

Which, as explained above, depends jointly on the positions of both pairs of nodes. To avoid this complication, we will occasionally work with a convenient upper bound on this optimal capacity:

$$C_{UBmax}(r, \theta) = \mathrm{Max}[C_{concurrent}(r, \theta), C_{multiplexing}(r, \theta)]$$

In all cases, the expected throughput is found by integrating over the area of the R_{max}-radius circle around the sender:

$$\langle C_i \rangle (R_{max}, D) = 1/\pi R_{max}^2 \cdot \int_0^{R_{max}} \int_0^{2\pi} C_i(r, \theta)\, r \cdot d\theta dr$$

In most cases, this integral cannot be evaluated analytically, but we can investigate it numerically in the physical regimes of interest.

3.2.3. The capacity landscape

To begin our investigation, what does a wireless channel look like under contention according to our model? Before we average away the details, how can we visualize the link capacity "landscape"? Some basic intuitions can be gathered from the style of plot shown in Figure 2. Here, we show link capacity as a function of receiver position – a capacity map. Each plot shows $C_i(r, \theta)$ for a particular choice of interferer distance D and concurrency or multiplexing; these plots include no averaging.[6]

[6] For clarity, in these plots we ignore shadowing, setting σ to zero.

At the transmitter's position, the origin, capacity rises to a tall, narrow peak (theoretically unbounded according to the model, though this is of little practical significance), and variously falls off elsewhere. Under multiplexing, capacity is independent of the location of the interferer. It slopes down smoothly with the receiver's distance from the transmitter, everywhere providing half the capacity of the contention-free channel. Under concurrency, however, capacity depends on interferer location. The three concurrency plots illustrate capacity at three different values of interferer distance D. In each case, a "hole" in the network appears around the interferer (visible in the plots as a dimple on the x-axis, and also as the white regions of Figure 3 below), but more broadly, capacity throughout the landscape trends downward as the interferer approaches the center.

Note that this is rather unlike the fixed-range, "cookie cutter" model of interference. When the interferer is far away, overall capacity remains largely intact under concurrency. As the interferer approaches, capacity everywhere begins to suffer, first slowly, then rapidly, but never abruptly. In the limit, when the two transmitters are coincident (not shown), capacity isn't quite zero, but it is extremely poor compared to taking turns – no receiver has an SNR better than 0dB.

3.2.4. Carrier sense behavior: a first look

What kind of MAC behavior can we expect from two competing sender-receiver pairs, according to our model? There are two easy limiting cases. First, when the pairs are very far apart ($D >> R_{max}$), concurrency is clearly optimal – for all possible receiver locations (within the network range), the pairs have little mutual interfering effect. Carrier sense handles this properly, since sufficiently distant senders will have mutually sensed power less than any reasonable sense threshold. Second, when the pairs are very close together ($D << R_{max}$), unless the signal powers are extremely low (generally below the noise floor), multiplexing is nearly always optimal – except for a tiny handful of receiver locations sitting practically on top of the transmitter, concurrent transmission would lead to an SNR approaching 0dB, regardless of receiver position. For sufficiently close separations, carrier sense also makes the right choice here. The crucial element in both these limiting cases is that, for a given D and R_{max}, essentially all possible receiver configurations are in agreement – either they all prefer concurrency or they all prefer multiplexing. So long as all receivers are in agreement, with appropriate choice of threshold,

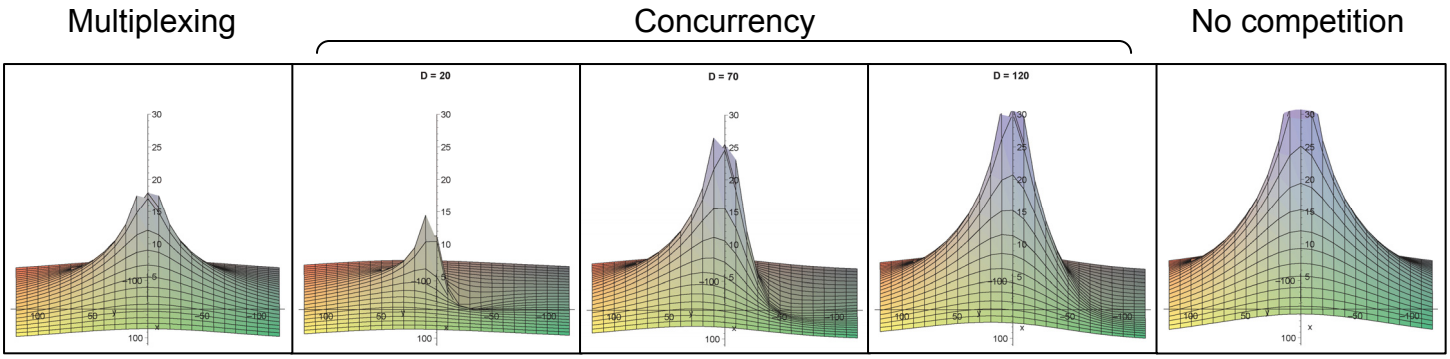

Figure 2 – Capacity "landscape" $C_i(r, \theta)$; $\alpha = 3$, $\sigma = 0$, $P_0 / N_0 = 65$dB. Depicts capacity as a function of receiver position, with the sender at the origin and an interferer on the x-axis at distance D. D is identified explicitly for the concurrency cases, but not for multiplexing, where

carrier sense can be made optimal in the low-shadowing limit.

As results in subsequent sections show (e.g. Figure 4), for reasonable parameter regimes, these "near" and "far" limiting cases or close approximations to them actually cover much of the relevant range of D. In such regions, carrier sense behaves nearly optimally. The "transition region" in between the two limits, however, necessitates compromise between the wishes of potential receivers who would prefer concurrency and those who would prefer multiplexing; no one threshold can satisfy them all. Figure 3 illustrates the geographical distribution of receiver preferences, for the same model scenario as Figure 2; white and light grey identify a preference for multiplexing, while dark grey marks preference for concurrency. Observe that, for a nearby interferer at D = 20, a single choice, multiplexing, is optimal for all R_{max} up to about 100. Similarly, for a distant interferer at D = 120, pure concurrency is optimal for all R_{max} up to about 50. In between, however, at D = 55, receivers are split nearly down the middle, half preferring concurrency and half preferring multiplexing. Limited to a single decision threshold regardless of the diversity among receivers, carrier sense efficiency must fall below optimal. One cannot immediately draw quantitative conclusions, because this illustration does not quantify the magnitude of the throughput sacrificed, but it does illustrate the basic dilemma.

Besides the distance between transmitters, another crucial factor in MAC behavior is the effective size of a network. In "long range" networks (R_{max} large), links are weak and network capacity tends to be dominated by the noise floor (thermal noise), even in the presence of interference. Receivers close to the interferer see significant interference, but most receivers are farther away, where the interference is weak and blends into the noise floor; the effects of interference become increasingly localized to the vicinity of the interferer. This represents a more difficult case for carrier sense, because receivers close to the interferer strongly prefer multiplexing while the rest prefer concurrency – as before, a problem of agreement. On the other hand, in "short range" networks, an encroaching sender's interference gradually smothers the entire network, becoming far too potent to ignore before any differences within the network are a concern. Once interference is strong enough to matter somewhere, its effects are felt everywhere, because it is already too powerful for path loss to drive it into the noise floor within the network range. Not surprisingly, carrier sense performs extremely well applied to this case.

3.2.5. Performance results

We're now in a position to look at the quantitative results of the model. The tables below report carrier sense throughput as a percentage of optimal MAC throughput, computed in Maple with

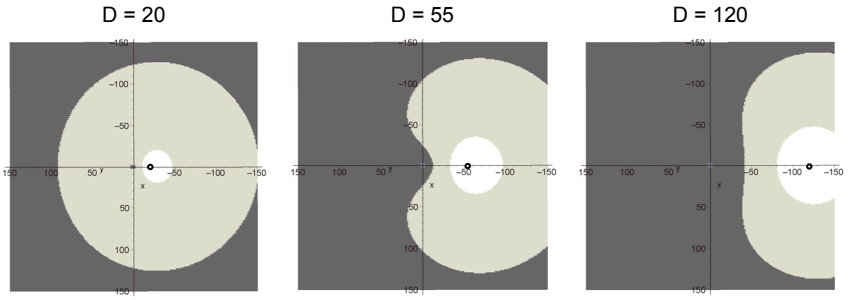

Figure 3 – Receiver preference regions: a receiver in the dark shaded areas prefers concurrency, in the light shaded areas prefers multiplexing, and in the white areas prefers multiplexing and will be starved (<10% of C_{UBmax}) without it. Circle marks

Monte Carlo integration, across a representative set of points in the model's parameter space: short and long range R_{max} (within the bounds of typical WLAN capabilities), near, transition, and far interferer distance D, and values of α and σ around typical real-world values. We have two important questions to answer: Is carrier sense performance adequate even when necessarily sub-optimal, namely, in the transition region and at long range? And does that performance hold up without specially tuning the sense threshold for each different environment?

The first table shows efficiency assuming a fixed $D_{threshold} = 55$, under fixed $\alpha = 3$, $\sigma = 8$dB:

Rmax \ D	***20***	***55***	***120***
20	96%	88%	96%
40	96%	87%	96%
120	89%	83%	92%

Carrier sense performance is extremely good overall, drooping slightly in the transition region and at long range. For comparison, the next table shows the same configurations but with thresholds optimized by the criteria in section 3.3.3 for each scenario:[7]

Rmax \ D	***20***	***55***	***120***
20 ***($D_{thresh} = 40$)***	93%	91%	99%
40 ***($D_{thresh} = 55$)***	96%	87%	96%
120 ***($D_{thresh} = 60$)***	89%	83%	92%

Very little change is observed, showing that carrier sense under such configurations is quite robust to small variation in threshold (or environment). We omit figures showing alpha varying from 2 to 4 and sigma from 4db to 12dB, but again, very little change is observed.

3.3. A deeper look: the simplified model

Why does carrier sense perform so well, and why is it robust to threshold variation? In this section, we explore these questions in detail using a simplified model: we eliminate random shadowing by setting $\sigma = 0$. Afterward, in section 3.4, we will reintroduce shadowing.

[7] The performance numbers are not strictly better because shadowing eliminates the notion of a globally unique optimal threshold; a trade-off exists between conservatively favoring small D and aggressively favoring large D. See section 3.4.

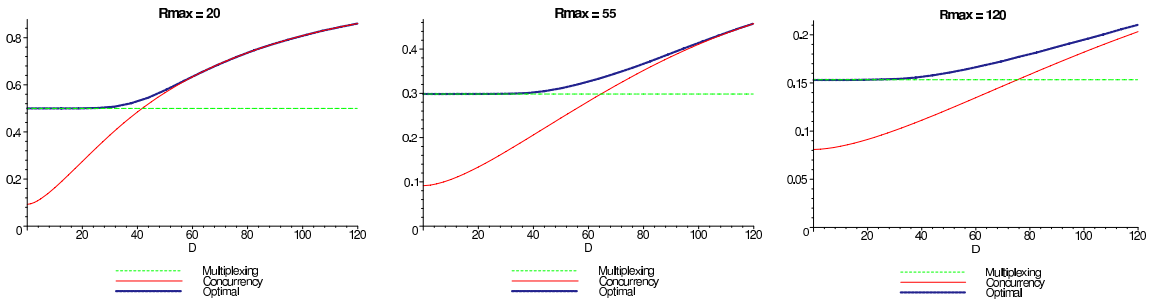

Figure 4 – Average MAC throughput curves for the non-shadowing model. $\alpha = 3$, $P_0 / N_0 = 65$dB.

3.3.1. Average throughput

To explore in more depth, we begin by plotting some relevant slices though the model's parameter space, showing how it behaves as a function of interference. In Figure 4, we plot our model's predicted throughput, averaged over several different radii R_{max}, against inter-sender distance D.[8] In each graph, the three curves represent capacity under three different MAC policies: multiplexing ($C_{multiplexing}$), concurrent transmission ($C_{concurrent}$), and the optimal MAC policy (C_{max}), which perfectly decides between multiplexing and concurrency on a case-by-case basis, taking the positions of both receivers into account. The horizontal scale is in multiples of the 65dB distance, as in section 3.2.2, while the vertical scale is normalized as a fraction of $R_{max} = 20$, $D = \infty$ throughput. Multiplexing throughput is, of course, independent of inter-sender distance, while concurrency throughput varies greatly, from little (though not zero) when the senders are coincident to ultimately twice the multiplexing throughput when the senders are widely separated. In this simplified (deterministic, non-shadowing) model, carrier sense throughput is a piecewise curve made up of the multiplexing curve to the left of the threshold $D_{threshold}$ and the concurrency curve to the right (see Figure 5). This is, in general, slightly less than the optimal throughput (C_{max}) curve. Also notice how optimal throughput approaches carrier sense throughput at both ends of the graph, converging to the concurrency branch as D grows large and the multiplexing branch as D becomes small.

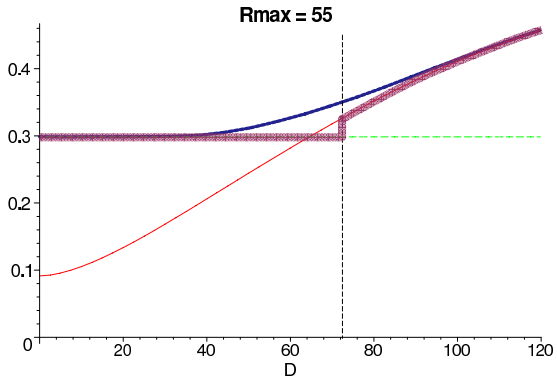

Figure 5 – Plot of $R_{max} = 55$ with carrier sense throughput for a particular chosen threshold (vertical line) highlighted.

[8] Note that, although they look somewhat similar, these graphs are not merely scaled versions of one another. To attain the same results at different length scales is possible but requires also scaling power with r^{α}.

As explained in the previous section, in both limits of D, essentially all possible receivers prefer the same choice, and so carrier sense performs all but optimally. On the other hand, in the transition region between the two extremes, different possible receivers disagree, those nearer the interferer preferring multiplexing and those farther preferring concurrency. The need to compromise on a single threshold for all receivers results in the visible gap between the best possible carrier sense throughput (Max[$<C_{multiplexing}>$, $<C_{concurrent}>$]) and optimal throughput in the vicinity of the crossover between the concurrency and multiplexing curves.

Although different receiver locations do disagree within the transition region, with bitrate adaptation, the magnitude of their disagreement is tempered. Situations commonly thought of as "hidden terminal" scenarios, where a node prefers multiplexing but doesn't get it, are not usually a disastrous case of "concurrency makes transmissions fail" but rather "a less-than-ideal bitrate is needed to succeed". Similarly, "exposed terminals" are not so much "concurrent transmissions would succeed" but rather "somewhat better throughput would be achieved with concurrency, though at a lower bitrate". Indeed, while there may be transient losses as the radio adapts, given properly chosen bitrates, the traditional notions of hidden and exposed terminals no longer really make sense. We could still refer to "hidden terminal inefficiency" (the gap between carrier sense and optimal throughput found to the right of the threshold, light shaded region in Figure 6) and "exposed terminal inefficiency" (the gap between carrier sense and optimal to the left of the threshold, dark shaded region in Figure 6), but aside from one exception discussed below, rarely are nodes truly "hidden" or "exposed".

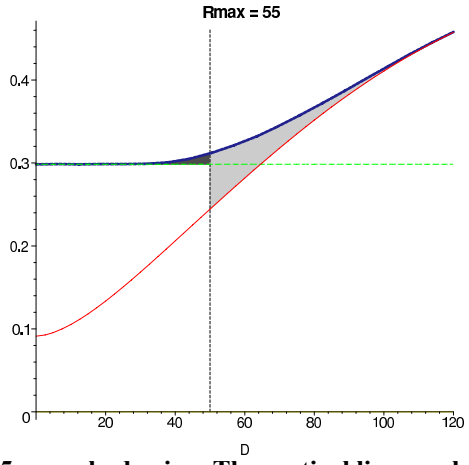

Figure 6 – Shaded plot of R_{max} = 55, non-shadowing. The vertical line marks the carrier sense threshold. Dark shading is inefficiency under multiplexing and light shading is inefficiency under concurrency, of which the light shaded "triangle" below the multiplexing line represents inefficiency due to poor threshold choice.

3.3.2. Discussion: Optimality

Sections 3.2.4 and 3.3.1 discussed why carrier sense approaches optimality in the limits of large and small D. But what about in the transition region? Why does carrier sense perform reasonably well there even though it necessarily falls short of optimal?

One of the most important reasons is the "non-locality" of interference – how the deleterious effects of interference are spread smoothly in space, rather than being localized to a cookie-cutter region. To take the extreme example, if interference caused the same fraction of capacity to be lost everywhere, increasing steadily as the interferer approached, there would be zero

disagreement among receivers; with a perfectly chosen threshold (and ignoring shadowing), carrier sense would be inherently optimal. On the other hand, at the opposite extreme, if interference were confined to a sharp-edged region within which concurrency was rendered strictly impossible (a flawed but conceptually convenient model of fixed-rate transmission), carrier sense would be distinctly less attractive. In the worst case (D just to the right of the threshold, the interferer just on the cusp of the transition to multiplexing), while many receivers would be unaffected, the rest would have no capacity at all.[9] In this extreme, neither concurrency nor multiplexing could possibly satisfy both groups of receivers simultaneously.

While interference in practical systems fits neither of the extreme caricatures above, it does show significant non-local, even global impact. The "landscape" plots of Figure 2 above provide a good illustration – comparing the concurrency plots against the "no competition" plot shows that the effects of interference are both widely spread and smoothly varying. Significantly, throughput worsens gradually for receivers closer and closer to the interferer. A fixed bitrate modulation, unable to survive at low SNR and unable to advantageously exploit high SNR, would transform this smooth SNR gradient into a step-like drop in throughput. Like the pessimistic example above, this becomes a dicey situation for carrier sense – no one threshold could satisfy receivers on both sides of the step. With adaptive bitrate, however, differences among receivers are not nearly so serious a concern. Some receivers might achieve less than optimal capacity, but by adjusting the transmitter's bitrate, they still generally get something reasonable.

Another important source of non-locality arises from geometry and dimensionality: in two and three dimensions, surface area grows rapidly with distance, meaning that many of the receivers, and hence much of the impact of interference, are found in the weakly interfered "fuzzy edge" rather than in the central, heavily interfered region (represented by the white regions in Figure 3). Together, the net result of these non-local effects is that while many receivers may still prefer concurrency within the transition region, enough interference reaches them that multiplexing isn't such a bad alternative, and vice versa. The number of severely impacted receivers remains small.

As alluded to in section 3.2.4, however, the degree of locality does depend on the size of the network and the distance to the interferer. Noise-dominated, long-range networks, where interference tends to decay to insignificance before reaching the far side of the network, experience much more strongly localized interference effects than short-range networks. In short range networks, where interference is unimportant until an interferer is substantially closer to the center of the network (in absolute terms), harmful interference is too strong to decay into the noise floor. Moreover, such a distance, while closer in absolute terms, remains farther relative to the size of the network, meaning that the interference level varies significantly less across the network than in the long range case. By the time an interferer is close enough to produce a substantial differential impact across the network, the network has long since switched to multiplexing. Consequently, with strong non-locality and little receiver disagreement, carrier sense is significantly more efficient in short range networks. This difference also has a striking

[9] Indeed, the threshold optimization method of section 3.3.3, which yields maximum average throughput, would leave roughly half the receivers to starve.

impact on the behavior of optimal thresholds, with implications for fairness, which we discuss in the next section.

Another important reason for carrier sense efficiency is the constraint that any optimal MAC must compromise over both pairs of nodes. Carrier sense with a reasonable threshold always picks a compromise favoring the "common case": most receivers have intermediate SNRs, and carrier sense provides such typical receivers with good performance (true by construction for a well-selected threshold). Receivers with unusual SNRs (e.g., those in a deep shadow or very close to the sender or interferer) tend to dislike some decisions made by carrier sense, but they are in the minority. A more sophisticated MAC might hope to improve upon carrier sense throughput by exploiting these "uncommon case" configurations, but to do so, it must not severely disadvantage the competing sender-receiver pair, or the result will fail to provide any aggregate improvement. However, most competing pairs *will* have typical SNRs and *will* dislike such aggressive exploitation; after all, the whole point was to make an unusual decision for an *uncommon* node configuration. For this to be a grand, overall win, *both* pairs of nodes must be in an uncommon configuration. This significantly weakens the potential gains from an aggressive MAC relative to a common case compromise MAC like carrier sense.[10]

3.3.3. Picking a threshold

What choice of threshold is best for carrier sense efficiency? In the non-shadowing case, we can pick a threshold that simultaneously minimizes average inefficiency for all D. Fairness and the distribution of allocations do vary as a function of threshold, but in terms of average throughput alone, the optimal threshold is the D at which the concurrency and multiplexing throughput curves cross, the point where concurrency provides half of the competition-free capacity. This fact is illustrated graphically in Figure 6 above, which depicts the inefficiencies associated with a suboptimal threshold. Whenever the threshold is leftward of optimal, a "triangle" region of added hidden terminal inefficiency appears (the portion of the light shaded region below the multiplexing line). Similarly, when the threshold is rightward of optimal, a triangle region of added exposed terminal inefficiency appears. These are both eliminated at the concurrency-multiplexing intersection point, which is therefore the best possible choice.

One can thus analytically determine the optimal threshold by solving for the intersection of the concurrency and multiplexing curves. However, this relies on knowing the value of R_{max} and the particular parameters of the propagation environment. That is unreasonable when a default value needs to be programmed in at the factory. Fortunately, as we discuss in 3.3.4 below, the precise choice of threshold turns out not to matter too much. As long as the threshold is somewhere in the transition region, the additional losses due to a suboptimal threshold will be small. So, our strategy can be to split the difference, picking a threshold roughly in the middle of the span of optimal thresholds for the typical operating range of our hardware. For example, given our usual $\alpha = 3$, 802.11g's bitrate flexibility ranges from around $r = 20$ to $r = 120$ (as explained in section 3.2.2). $R_{max} = 20$ corresponds to an optimal threshold about $D_{thresh} \approx 40$, and $R_{max} = 120$ corresponds to $D_{thresh} \approx 75$. So, a $D_{thresh} \approx 55$ (equivalent to $P_{thresh} \approx 13$dB) would be a reasonable compromise.

[10] Formally speaking, this represents the gap between $<C_{max}>$ and $<C_{UBmax}>$.

Now, resuming the discussion of short versus long range, observe that, in our $R_{max} = 20$ example, the optimal threshold is roughly twice the network range, a demonstration of strong non-locality. In general, a short range network has an optimal threshold whose equivalent distance is well outside the network boundaries. On the other hand, a long-range network (e.g. $R_{max} = 120$) may have an optimal threshold *inside* the network, waiting till the interferer approaches close enough to impact the network in all directions.[11]

As a convenient consequence, we can use the position of the optimal threshold relative to the network boundary as a quantitative distinction between short and long range. $R_{thresh} < R_{max}$ is a good marker of the genuine long range regime, while $R_{thresh} > 2R_{max}$ is a reasonable, rough boundary for the beginning of true short range. With these criteria in mind, an interesting property emerges: In short range networks, not only is average throughput good, but every receiver has a reasonable share, because whenever concurrency is employed, interferers are too far from the network to have a localized impact. In long range networks, however, while throughput remains good for most receivers, when an interferer begins to tread inside the network range itself, a small, nearby fraction of receivers gets smothered in interference, receiving little throughput at all. The white regions in Figure 3, showing receivers that receive less than 10% of C_{UBmax} under concurrency, illustrate this effect. In the D=120 and D=55 frames, the white region has the potential to be inside a sufficiently long-range network that is still operating under concurrency; the receivers in this region could reasonably be described as hidden terminals. Overall, while average throughput in long range networks remains robust, fairness can suffer some.

3.3.4. Discussion: Threshold robustness

The previous section addressed the question of why carrier sense has good performance when pre-configured with a threshold optimal for the propagation environment and network size. But why does performance remain good even in the absence of such fine-tuning?

It's not hard to see why carrier sense is robust at least to small variations in threshold. In order for small variations to produce bad behavior, the gaps between concurrency and multiplexing throughput (as in, e.g., Figure 4) would need to abruptly grow very large on either side of the optimal threshold – where the gap necessarily vanishes. But, as the model demonstrates (e.g. Figure 4), interference does not behave in such an abrupt manner. Given adequate bitrate adaptation, the network-wide impacts of interference are quite gradual with interferer distance, changing only on the length scale of the network radius.[12]

But what about large-scale variations in the parameters of the model itself? Why don't variations in α, σ, and R_{max} make the optimal threshold swing so broadly that a custom-tuned threshold is simply a necessity?

---

[11] Indeed, there is a third regime beyond "long range" that we do not consider in this thesis, where concurrency is *unconditionally* optimal – multiplexing is never preferred. This "extreme long range" regime is the realm of CDMA systems, where mechanisms like carrier sense merely get in the way. In this regime, however, capacity per user tends to be extremely poor, and efficiency is attained only for very large numbers of concurrent transmitters. As a result, data networking hardware is seldom targeted for this regime.

[12] Variations of this statement can be proven mathematically. E.g., for $\alpha = 3$, $\sigma = 0$, the slope of the concurrency curve (in our $R_{max} = 20$ normalized capacity units) is bounded above by $1.37 / R_{max}$ for all $D > R_{max}$.

The answer appears to arise in large part from the "sweet spot" operating regime chosen by the vast majority of data networking hardware. While often capable of ranging a bit wider, most data networking hardware is designed to operate in the regime roughly around 10-25dB SNR; 802.11 [IEEE07], Bluetooth, and 802.15.4 [IEEE06] all target this region. This is likely a practical consequence of capacity and economic considerations. It's valuable to have an SNR well above the noise floor, in order to provide a reasonable amount of capacity and a reasonably affordable receiver, but increasing the SNR further leads to diminishing returns as the logarithm in Shannon capacity kicks in, and doing so often runs up against regulatory and power handling constraints as well.

Conveniently for carrier sense, this operating regime also turns out to have special significance for MAC behavior: it is the intermediate region between long range and short range. This can be demonstrated using the quantitative criteria proposed in section 3.3.3, identifying long range networks as those with optimal thresholds inside the network boundary and short range networks as those with optimal thresholds substantially outside the network boundary.

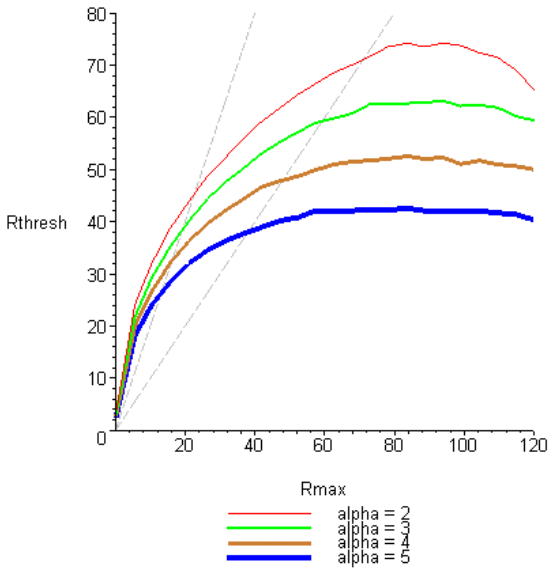

Figure 7 - Optimal threshold (expressed as the equivalent distance at α = 3, for consistency) versus network radius for several values of α, with σ = 8db (included because shadowing has a significant qualitative impact at long range). The erratic ripples on the right are artifacts of the numerical solution method. Dashed lines mark $R_{max} = R_{thresh}$ and $R_{max} = 2*R_{thresh}$.

Figure 7 shows the variation of optimal threshold with the size of the network, tabulated at several different values of α, representing different propagation environments (α = 2 - 4 being typical). On the left lies the short range limiting behavior, thresholds scaling towards zero roughly as the square root of R_{max}[13] and clustered closely together in spite of α variation. On the right lies the long range limiting behavior, where threshold growth tapers off in R_{max} but spreads out in α. In between, centered around R_{max} = 40 or so (corresponding to ~17dB SNR at the edge of the network), the curves show a gradual change in behavior from the one extreme to the other.

[13] Optimal $D_{threshold} \approx e^{-1/4} R_{max}^{1/2} N^{-1/2\alpha}$ (actual distance units, not α = 3 equivalents) for very short range networks, which can be derived by taking the limit as $N \rightarrow 0$ and approximating Δr as $D_{threshold}$.

The figure also plots the lines $R_{thresh} = R_{max}$ and $R_{thresh} = 2*R_{max}$ (dotted lines). The points at which these lines intersect the optimal threshold curves correspond to our boundary criteria for long range and short range, respectively, for each plotted α. These boundary lines neatly enclose the behavioral changes in the threshold curves, showing that the changes do in fact correspond to the crossover from short range to long. For typical $\alpha \approx 3$, this range is roughly $18 < R_{max} < 60$, equivalent to 12dB < SNR < 27dB at the edge of the network.

Why does the fact that data networking hardware favors the intermediate regime help carrier sense? In the limit of short ranges, optimal threshold scales rapidly with network size; in this limit, while carrier sense performs well for a well-tuned threshold, thresholds are not particularly robust to uncertainty in power or distance. On the other hand, thresholds vary quite slowly in the long range limit, owing to the locality of interference. However, as we discussed in section 3.2.4, carrier sense efficiency falls off; optimal thresholds are robust but still perform poorly. Further, given the wide variation with α in the long range, thresholds there are not particularly robust to varying propagation environments. In the middle, however, is a compromise that approaches the best of both worlds – good performance *and* robust thresholds. For completely unrelated reasons, this is the primary operating regime for wireless networking gear.

3.4. Re-introducing shadowing

In this section, we return shadowing to our model. Real environments are full of arbitrarily arranged obstacles and reflections, and one might expect these hot spots and shadows to harbor scores of hidden and exposed terminal configurations. However, we show that such irregularities do not ruin the intuitions of the previous section and that carrier sense continues to provide close-to-optimal average performance.

One might imagine, for example, that a hidden terminal configuration could be constructed by inserting barriers into the environment as in Figure 8. Radio propagation, however, is not so easily confined. Most building materials are not particularly opaque to radio; typical attenuation through an interior wall is less than 10dB [COST231 §4.6-4.7]. However, even if the obstruction were opaque (for example, a metal barrier), a substantial reflection from the far wall would connect the two transmitters. Reflections are exceedingly difficult to suppress; typical reflection losses are less than 10dB. Yet, even if there were no far wall, only open space, a weak signal would still round the corner, as a result of diffraction. Using the knife-edge approximation and a 5-meter distance to the barrier, the diffraction loss at 2.4GHz would be around 30dB.

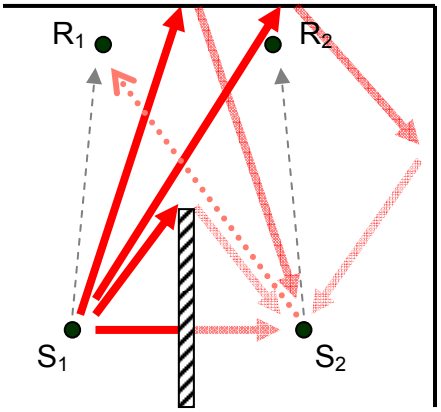

Figure 8 - Conceptual graphic of propagation pathways past a barrier. The dashed arrows show desired transmissions, while the dotted arrow shows a potential source of "hidden terminal" interference. The thick, red arrows show several propagation paths for a carrier sense signal.

Radio propagation thus becomes quite complex and difficult to constrain – so complex, in fact, that one can reasonably apply the central limit theorem and lump together all the effects into a single Gaussian random variable. This is the origin of lognormal shadowing [Akaiwa97]. Empirically, the standard deviation is typically of the same order of magnitude as the losses quoted above, typically 4-12dB (see section 2). Such uncertainty, though significant, is not sufficient to destroy the efficiency of carrier sense.

This lognormal variation affects carrier sense by adding randomness along three dimensions: the signal power at the receiver, the interference power at the receiver, and the interferer's power at the transmitter, used in making carrier sense decisions. No longer defined uniquely by distance, each of these powers now has an independent random component.[14] To estimate (somewhat pessimistically) the potential effect on a sender's ability to estimate its receiver's SNR, the three effects can be summed, yielding $\sigma_{SNRest} = \sigma\sqrt{3} \approx 14$dB uncertainty (less if the interferer power is comparable to the noise floor), assuming $\sigma = 8$dB shadowing as below.

To help visualize the implications, we can consider what happens under roughly equivalent distance variations lumped onto a single, appropriate model parameter such as interferer distance D. Under $\alpha = 3$, 14dB's equivalent in path loss is a distance factor of about 3x. For the most highly impacted receivers, when approximately lumping all three shadowing effects into D, a value closer to 2x turns out to be more appropriate.[15] Although this variation goes both positive and negative, only one direction contributes to carrier sense inefficiency, because only error in one direction can make any particular choice in the binary "concurrency-vs-multiplexing" decision undesirable. Similarly, only variation for receivers whose preferred choice is near the decision threshold contributes to inefficiency, and for these nodes, the impact of carrier sense's decision is already limited. While 2-3x is significant, since the distance to an interferer can range over much more than a factor of three, in the limiting cases, receivers remain reasonably

[14] For this analysis and throughout, we assume that the shadowing distributions are uncorrelated. This is not quite true, but it is good enough for our purposes.

[15] For receivers approximately equidistant between sender and receiver, around 3/2 x variation in D (equivalent to 6.5dB) covers 8dB interference power variation. Sense power uncertainty contributes 8dB directly, and effects roughly equivalent to sender power variation require just shy of 4dB. Combining these gives about 11dB, roughly 2.3x distance.

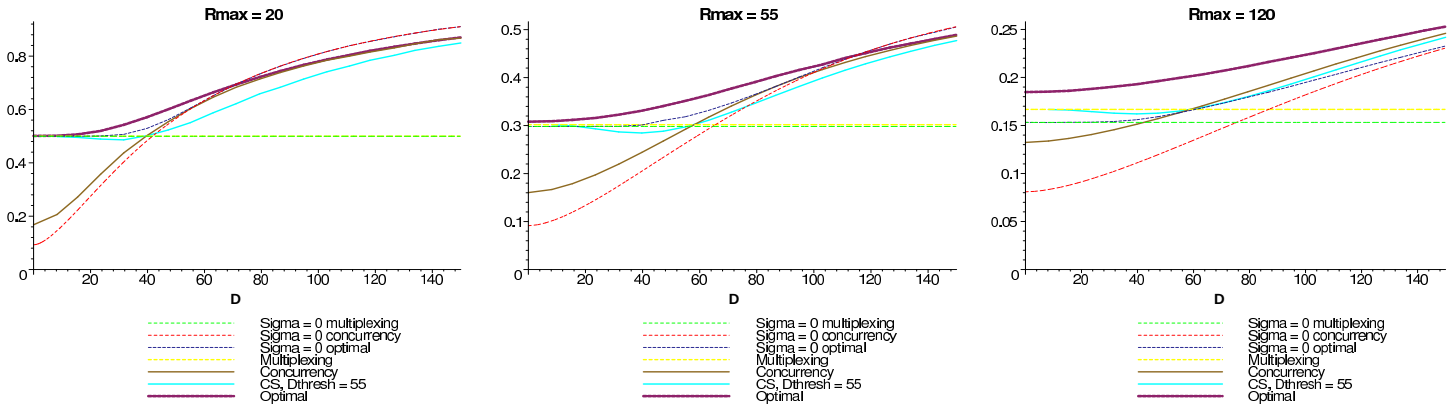

Figure 9 – Average MAC throughput curves for the model with 8dB shadowing, along with non-shadowing curves for reference.

homogeneous and carrier sense remains reasonably predictive. On the other hand, while the intermediate cases do become more difficult, the precise choices made by carrier sense there remain of lesser importance.

To take a concrete example, in a short range network of size $R_{max} = 20$ with threshold $D_{thresh} = 40$ (close to the $\sigma = 0$ optimum), an interferer that, to the receiver appeared to be at D = 20, would have about a 20% chance of appearing to the sender as beyond D_{thresh}, thereby triggering concurrent transmission. This mistake would leave the receiver with a very low, sub-0dB SNR about 20% of the time (approximately the fraction of the R_{max} disc's area closer to D = 20 than to the sender). Combining the probabilities, we have that for the given interferer position, the effects of shadowing on carrier sense would cause very poor SNR in around 4% of configurations but otherwise would behave reasonably most of the time.

Figure 9 shows the detailed, quantitative results of our shadowing model, using $\sigma = 8$dB and the same $\alpha = 3$ path loss as before. The non-shadowing results from Figure 4 are also plotted on the same axes, for comparison. The results are qualitatively very similar. Short range carrier sense does extremely well, hugging the optimal curve closely, while long range is somewhat less efficient. Best performance is found in the limits of near and far interferer distance, while some efficiency is lost in the middle. However, a few differences are notable:

First, because sensed interferer power has a random component, the sender does not always make the ideal decision as a function of interferer distance. Mistakes lead to a bit of performance degradation, and for a given interferer distance, expected throughput becomes a weighted average between concurrency and multiplexing, rather than a binary choice.[16] Figure 9 reflects this in the plotted carrier sense throughput, which interpolates smoothly between concurrency and multiplexing, always hanging slightly below the ideal piecewise throughput. This effect also causes the efficiency degradation of the transition-region to cover a wider range of D.

[16] A notable consequence is that there is no longer a uniquely defined "optimal" threshold – while the old, $\sigma = 0$ optimum remains a good choice, a tighter threshold will slightly improve expected performance for large D by reducing the chances that the interferer will spuriously appear to be closer than the threshold, in exchange for worse performance at small D, and vice versa.

Second, the unpredictable variation in signal powers adds some random variation (and hence "disagreement") among receivers in their preference between concurrency and multiplexing; shadowing makes it harder to pick a single choice that satisfies all possible receivers. This tends to increase the gap between carrier sense and optimal throughput, as Figure 9 illustrates.

Finally, the nonlinearity of capacity as a function of dB SNR means that incorporating zero-mean variation to the model actually has a net positive impact on average capacity. This is particularly significant in long range networks and under concurrency, where the effect can be summarized as "you can't make a bad link worse than no link, but you *can* make it a whole lot better". Under shadowing, the low SNR regime is inherently unreliable – some nodes will have so little capacity that they are all but disconnected, while other nodes will have surprisingly good links – and the average is higher than without shadowing. As a result, in the long range, concurrency fares surprisingly well, albeit at the expense of worsening the already poor fairness of long range networks. The net effect on average throughput is a reduction in the gap between concurrency and multiplexing in long range networks and a leftward shift in their optimal thresholds, both visible in the D=120 frame of Figure 9.

This reduction and shift have the interesting consequence of making bad carrier sense decisions less frequent and less costly, mitigating some of the above negative effects of shadowing on capacity. It also further reduces the need for threshold tuning; systematic trends among nodes tend to get blurred out anyway. Finally, it serves as a word of caution to be wary of worst case analysis in MAC research: a MAC-layer worst case is only a worst case if it's worse than the failures of the PHY.

4. Experimental results

This section presents experimental results from a large indoor testbed, comparing the qualitative predictions of the theory with behavior from contemporary hardware. We find mainly broad agreement, as well as a few interesting discrepancies.

Our experiments measure average throughput for competing sender-receiver pairs under concurrency, multiplexing, and carrier sense. We run our tests on an indoor testbed constructed from mixed Atheros AR5212 and AR5213-based 802.11a/b/g wireless adapters equipped with one "rubber duck"-style antenna each, installed in roughly 50 Soekris single-board computers scattered about two closely-coupled floors of a large, modern office building. All experiments run in 11a mode.

To measure throughput, each of the two senders attempts to send 1400-byte packets continuously for 15 seconds, and we count the number of packets successfully received at the intended receiver. For concurrent throughput, we disable carrier sense and run all transmitters simultaneously. For multiplexing throughput, we run each sender-receiver pair alone, one after another. For carrier sense, we enable default hardware carrier sense and run both transmitters simultaneously. To determine throughput given an optimal bitrate, we repeat every run at each of

6, 9, 12, 18, and 24Mbps[17], independently identifying the maximum throughput bitrate for each transmitter.

We follow a practical variant of the usage assumptions of the model. Rather than communicating with nodes within a given geometric range, senders communicate with nodes within some link-level metric. In our case, we use packet delivery rate at 6Mbps as our metric. A situation where all receivers have very good delivery rates corresponds well to a short-range scenario. Intermediate delivery rates correspond to more of a long-range scenario. (We do not extensively probe poor delivery rates, because that necessarily pushes beyond the rate adaptability of the hardware when in 11a mode.) We break experiments into two categories, pairs with links better than 94% delivery at 6Mbps (resulting in an average SNR of about 27dB, which is roughly similar to that of an R_{max}=30 model network) and pairs with links between 80% and 95% (average SNR about 16dB, roughly similar to R_{max} = 70).

4.1. Short range

Figure 10 and Figure 11 below show the results of the short range (links 94% at 6Mbps) experiment. The first plots the performance of pure multiplexing and pure concurrency against hardware carrier sense performance (CS), as a competitive comparison. Each set of three vertically collinear points represents one run, one pair of transmitter-receiver pairs. The pink square shows total (combined) throughput in packets per second for concurrent transmission (CS disabled), the blue diamond shows total throughput for multiplexing, and the yellow triangle shows carrier sense throughput (on the identity line), for reference. In the absence of experimental variation and non-idealities, we would expect to see the each CS point coinciding with either its corresponding concurrency point or its corresponding multiplexing point. The preponderance of these points close to or below the CS identity line demonstrates that carrier sense performs quite well here.

The second figure plots the same data, this time against average RSSI between the two senders (measured separately), a network-level measure of sender-sender separation and the metric around which carrier sense is based. The vertical bars highlight points where CS performance is below optimal. At the very right (plotted in the column above zero, spread randomly for visibility) are the points for which no test packets were received, implying a signal strength somewhere below the receive threshold.

[17] Higher bitrates perform too poorly to be useful under our nonstandard, OpenHAL-based driver capable of disabling carrier sense, and 11g mode support, capable of lower bitrates, is buggy.

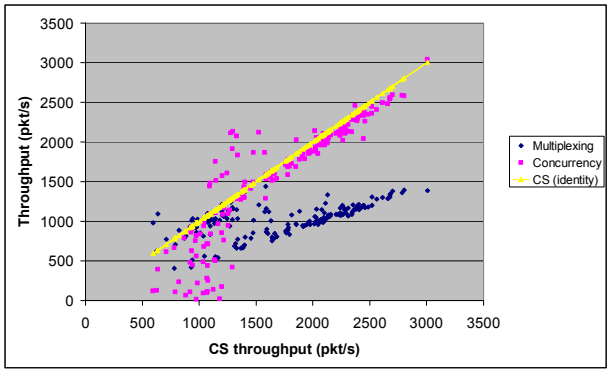

Figure 10 – Short range competitive comparison vs. CS.

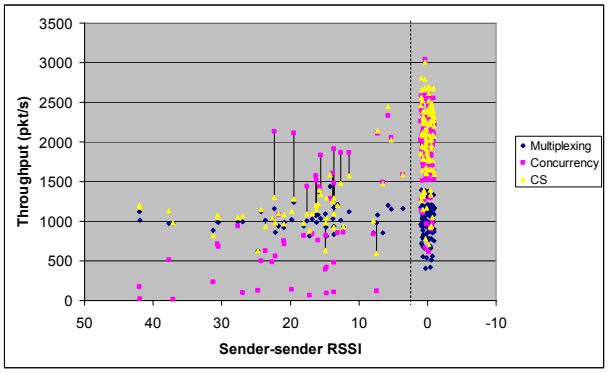

Figure 11 – Short range throughput vs. sender-sender RSSI.

Because link quality for all pairs is quite good, the main performance-limiting factor in these graphs is pair-pair interference, with which both x-axes are strongly correlated. The two plots show roughly the same set of features: a close-distance region on the left, where multiplexing and CS performance coincide and concurrency does quite poorly, a transition region in the middle (roughly 20dB-10dB in Figure 11), where concurrent performance catches up and sometimes exceeds both CS and multiplexing, and a long-distance region on the right, where CS and concurrent performance coincide and multiplexing lags behind by a factor approaching two. This breakdown is as the theory predicts.

An important feature here is that, in this short range data set, carrier sense is quite infrequently bested by multiplexing or concurrency. Even when it is, the gains are not especially compelling, limited to a small set of weakly exposed terminals. Averaging throughput over all runs (that is, over the slightly arbitrary ensemble of pairs and distances present in the testbed) gives the results in the table below. Carrier sense approaches the optimal strategy quite closely, consistent with theoretical predictions for very good behavior in the short-range case.

Optimal (max over strategies): 1753 packets / sec
Carrier Sense: 1703 pkt/s (97% opt)
Multiplexing: 1013 pkt/s (58% opt)
Concurrency: 1563 pkt/s (89% opt)

One notable oddity is that, even in cases of very wide pair-pair separation, CS performance often slightly exceeds concurrency. This conceivably might be due to time variation in the channel, properly exploited by carrier sense, causing concurrency to be suboptimal for a small fraction of the time. Alternatively, there may be some subtle experimental bias.

4.2. Long range

Figure 12 and Figure 13 below show the results of the long range (links 80% to 95% at 6Mbps) experiment. The format is the same as the short range plots, described above. This dataset is not quite as "long range" as we would like, but pushing farther into the long range regime runs up against the limits of bitrate adaptability in 11a mode, forcing many links to operate at the minimal 6Mbps even before interference. This both hurts concurrent performance and introduces behavior intermediate between variable bitrate and fixed bitrate, which we haven't attempted to model. Using 11g mode instead should reduce such difficulties in experimentally exploring

deeper long-range scenarios. Of course, there is always some adaptation floor, at which point the network becomes unreliable and our capacity model assumptions are violated.

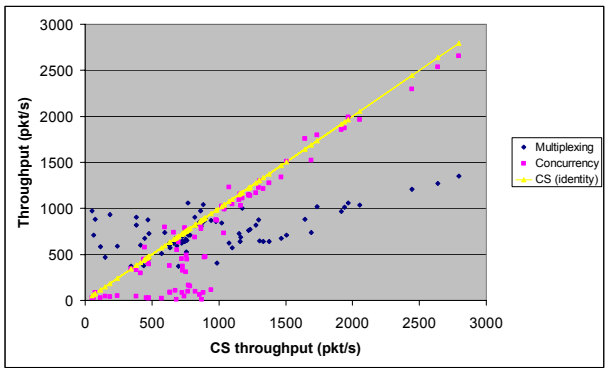

Figure 12 – Long range competitive comparison vs. CS.

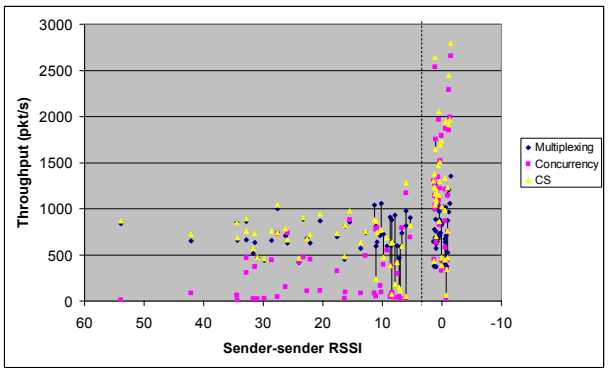

Figure 13 – Long range throughput vs. sender-sender RSSI.

Although not as clear-cut as the short-range case, this long-range dataset shows a similar set of features: a close-distance region where multiplexing is preferred, a transition region (roughly 10dB – 5dB in Figure 13), and a long-distance region where concurrency is preferred. In the plot against RSSI, these are cleanly visible, one after another from left to right. Note that the transition region mistakes consist mainly of undesirable concurrency – "hidden terminals" – rather than undesirable multiplexing, as the theory would predict given a threshold optimized for the "average case" rather than specifically for long range. Also note that the transition region, located just shy of 10dB, is several dB lower than the roughly 15dB of the short-range case. Again, we would expect this sort of shift from the theory.

In the competitive plot against CS throughput, however, the regions are somewhat muddled together. Many of the transition region's cases of undesirable concurrency experience a crash in throughput, and these cases end up on the left of the plot, rather than in the middle. This makes the graph harder to interpret.

One oddity about many of these "undesirable concurrency cases", however, is that they seem to be intermediate in throughput between pure concurrency and pure multiplexing. The corresponding pure concurrency cases crash badly, perhaps even more than the theory would assume (possibly due to limited bitrate flexibility or the fact that we used broadcast packets and did not have receive abort enabled, making it impossible to identify the desired packet at the MAC layer), but the CS cases fare somewhat better. One possible explanation is that the CS decision itself is fluttering due to noise, and so these cases aren't actually pure concurrency but instead are a mixture. Another possible explanation is asymmetric carrier sensing, where one node defers while the other does not, leading to a mix of concurrency and unfair multiplexing.

Although carrier sense in the long-range here is not quite as close to optimal as it was in the short-range (as the theory would predict), it is still quite good overall and significantly better than either pure multiplexing or pure concurrency. Averaging throughput over all runs, as before, gives the table below:

Optimal (max over strategies): 1029 packets / sec
Carrier Sense: 923 pkt/s (90% opt)
Multiplexing: 753 pkt/s (73% opt)
Concurrency: 709 pkt/s (69% opt)

4.3. Summary

With these experiments, we showed that our theoretical claims, and in particular, the good expected performance of carrier sense, are quite consistent with observed behaviors in commodity hardware. As long as we stayed reasonably within our hardware's bitrate adaptation range, our tests showed high average carrier sense performance, of the same order of magnitude as theoretical values. Short range performance was particularly good and free of starvation, as the theory predicted, while long range performance was solid on average, in spite of a limited number of cases where throughput fell significantly. We also saw carrier sense behavior splitting up as a function of interferer distance into three distinct regimes, near, intermediate, and far, just as the theory claims, with modest inefficiency in the intermediate regime but nearly optimal performance near and far. Thus, within the parameter range we explored, our experimental results concur with the theory, supporting its claim that carrier sense's average throughput leaves limited room for improvement.

5. Discussion and future work

Under reasonable assumptions, the preceding analysis and experimental results have shown that carrier sense has good overall performance, both in theory and in practice, solving the spatial reuse problem with an efficiency approaching optimal. With adaptive bitrate, there is no particular need for threshold tuning in WLAN-like scenarios. There are also no dramatic losses of efficiency due to differing channel conditions among the nodes, the differences that Karn first warned of in [Karn90]. Better exploitation of hidden and exposed terminals could improve behavior in several corner cases, but it would have little effect on average performance. Better treatment of hidden terminals, in particular, could improve fairness and reliability at long range – though long range links are still likely to be failure-prone without much deeper bitrate adaptation than commonly provided today.

Pursuing exposed terminals, however, seems significantly less interesting. Unlike hidden terminals, which, when they do happen to arise (i.e. § 3.3.3, 3.4), can be acute, localized failures, exposed terminals are merely diffuse opportunities for modest improvements in throughput. As our modeling and experiments showed, these improvements are typically *very* modest. The primary reason why our results are so much more pessimistic than earlier claims, for example, [Vutukuru08], is that we treat adaptive bitrate as a first-class consideration.

The goal of exploiting exposed terminals, better throughput through better concurrency, is deeply undercut by the use of bitrate adaptation. Both adaptive bitrate and increased concurrency are viable strategies to exploit excess channel capacity, but unless nodes are widely separated or SNRs are *extremely* low, adaptive bitrate is strictly more efficient. In informal experiments using the same test set-up as section 4.1 (short range, where exposed terminal scenarios are found), we observe that exploiting even the weak bitrate adaptation our driver supports (6, 9, 12, 18, and 24 Mbps) more than doubles average throughput compared to the base rate. The alternative,

perfectly exploiting the exposed terminals, provides just shy of 10% increased throughput.[18] Attempting to combine the two, exploiting exposed terminals on top of bitrate adaptation, yields only about 3% more than bitrate adaptation alone.

Our results suggest that the problem of wireless concurrency control merits a lightweight, "incremental" approach. As this thesis showed, a simple and widely used mechanism can handle the common case with high efficiency. The interesting challenges lie in improving the behavior of the corner cases. Because the corner cases are infrequent, it is fair to address them with blunt mechanisms that are wasteful of bandwidth, provided we only engage said mechanisms when necessary. For example, a variant of RTS/CTS might be an effective solution to the starvation problem in long range networks subject to intermediate-distance interferers. The crucial step, though, omitted from protocols like 802.11 and MACAW [Bhargavan94], is to include a (possibly heuristic) mechanism for detecting when RTS/CTS protection is necessary. Otherwise, not only is it a waste of bandwidth for its direct users, it is a waste of spatial reuse among their neighbors. If, instead, RTS/CTS were enabled only when, for example, a sender discovered that it was experiencing an extremely high loss rate to some receiver in spite of a high RSSI, one could expect nearly all the reliability benefits of RTS/CTS would accrue, yet without any significant cost to average network throughput.

There are also several interesting corner case problems related to the physical implementation of carrier sense. Common carrier sense implementations can exhibit several pathologies not captured by our theoretical model, such as threshold asymmetry (where only one node defers to the other), slot collisions (where two nodes that normally multiplex transmit concurrently due to randomly picking identical starting times from a limited pool of MAC-defined slots), and chain collisions (where nodes transmit after failing to detect packets whose preamble signatures were hidden beneath another transmission, thereby hiding their own preambles and perpetuating the pathological state; interestingly, this is particularly likely to strike research protocols that send long, uninterrupted bursts of packets without per-packet acknowledgments). Other interesting issues include carrier sense compatibility across unrelated PHY schemes (generally lacking at present) and challenges associated with convergent flows to a single node, where signals may be weak but the concurrency decision must be based not on carrier sensing but on the contents of the demodulated MAC header[19]. We have several ideas for possible future research to address these issues, if demonstrated to warrant attention.

The theoretical techniques demonstrated in this thesis are more broadly applicable than just to carrier sense. The essential idea, combining a function that roughly models adaptive bitrate capacity with a statistical model of radio propagation and aggregating the results over a simple spatial footprint, can be applied to a variety of other problems in wireless networking involving small numbers of nodes. For example, interference cancellation, which recently has attracted a lot of attention in the networking community, could likely be evaluated with these techniques,

[18] The results [Vutukuru08] reported were higher by a small constant multipliers. This appears to be because they include results only for high levels of concurrency; as they show, the potential gains of exposed terminal exploitation increase with increasing concurrency. Their best result, 47% average improvement, required *six* concurrent senders – and [Cheng06] suggests this level of concurrency is quite a rarity in typical WLAN deployments.

[19] An issue which 802.11 particularly exacerbates by transmitting the MAC header at full data bitrate rather than at base rate like the PLCP header.

providing a more abstract but much broader answer to the question, "is it useful?", than can be drawn from simple demonstrations with software radio. Such theoretical analysis would be a valuable complement.

6. Related work

Carrier sense has a long history. It has been in wide use in wired networks since at least [Metcalfe76] and has been applied to wireless networks in various modified, sometimes quite complicated forms, e.g. [Tobagi75] [Bhargavan94] [Fullmer95] [Rao05] [Vutukuru08]. Early efforts tended to focus around the hidden terminal problem and the inability of nodes to detect certain nearby transmissions. These difficulties were likely due to the primitive nature of the radio hardware at the time, however, with limited signal processing power and largely using fixed-bitrate, narrowband (and hence fading-prone) modulations. [Bhargavan94], in particular, not only used fixed-bitrate modulation but also a highly unusual "near-field" radio technology that, although avoiding fading, had an extraordinarily high path loss exponent that made the hidden terminal problem very real. Their MACAW WLAN proposal employed an astonishing 5-packet exchange for each frame (RTS-CTS-DS-DATA-ACK), in contrast to the two-packet exchange (DATA-ACK) that has become widespread practice on today's radios. Modern radio network specifications employ far-field radios, along with fading-resistant wideband modulations, largely protecting them from the capriciousness of radio's wave interference patterns, and most either support multiple bitrates or are presently being extended to add them.

In spite of criticisms of carrier sense such as [Jamieson05], recent proposed alternatives have gained very little ground in overall throughput. [Jamieson05], in particular, suggested that bitrate-dependent tuning might be a route to improving carrier sense performance, noting the very high "exposed terminal" concurrencies possible at low bitrates on high SNR links. However, using such low bitrates on those links already wastes many times over any possible gains due to concurrency; no reasonable bitrate adaptation algorithm should select them. In [Rao05], by scheduling transmission attempts, Overlay MAC Layer demonstrated impressive fairness properties but did not improve throughput. [Vutukuru08] claimed improved throughput with their contention graph-based MAC protocol, but only in the restricted case of fixed, low bitrate and very high concurrency ($n \geq 3$), on a network capable of supporting much higher bitrates.

For some network structures, however, the assumptions of the model presented in this thesis are unjustified, and the conclusions do not necessarily hold. For example, [Patra07] uses highly directional antennas for long, dedicated links spanning tens of kilometers. For such networks, non-carrier-sense MAC protocols may well be much more valuable. Long (three or more hops), over-saturated multi-hop paths are also out of scope, because they violate the independent receiver distribution assumption. GSM-style cellular networks are another example. By virtue of using full-duplex radios with non-overlapping frequency bands, they *cannot* implement carrier sense – but even if they could, the extreme asymmetry between handset and neighboring base station, legacy considerations, and a variety of other differences could make TDMA more attractive, already very well-suited to predictable, inelastic flows such as voice.

Information theoretically, it is possible to achieve higher throughputs than our model here allows [Gallager85], using physical layer techniques such as joint decoding and interference cancellation [Verdu98]. However, while such techniques have appeared to some extent in

CDMA cellular systems, they have seen very little commercial application in data networks, likely due to their increased hardware cost and limited benefit under low concurrency. It is not obvious whether meaningful additional capacity is available in practical data networking scenarios – though one might be able use techniques like those of this thesis to conduct an analysis.

An independent thread of research that also studies carrier sense with analytical models and often Shannon capacity is represented in, e.g., [Yang05]. These works tend to focus on worst-case analysis and maximum concurrency, which leads to interesting performance limits but doesn't directly pertain to the average case behaviors of realistic data networking scenarios, at which this thesis aims. They also tend to use less realistic radio propagation models that lack shadowing and regularly drop the noise floor term, which completely wipes the long range regime from view.

Other related analyses, e.g. [Zhu04], identify possible ways to improve on carrier sense in the long-chain multi-hop case. Such claims are not inconsistent with this work, as we have explicitly excluded multi-hop chains beyond two hops. Indeed, there are reasons to believe carrier sense should be approached quite differently for fully saturated, long-haul, multi-hop trunks lacking adequate congestion control.

Although we briefly mentioned practical hardware implementation issues for carrier sense in Section 5, there are several interesting questions here that are beyond the scope of this thesis. [Ramachandran07] provides a clear introduction to hardware carrier sense techniques and their sensitivities, as well as some power consumption implications. Some useful background on signal-level implementations and a present-day instance of a recurring compatibility challenge (old radios unable to carrier sense new packet formats, due to the use of preamble detection rather than energy detection) can be found in [Aoki06]. [Rahul08] demonstrates a novel albeit complicated approach to such compatibility challenges in the context wideband OFDM and argues against the practicality of the energy detection solution, at least in the multi-channel case. Threshold asymmetry has been experimentally observed several times, such as in [Rao05], and it showed up in our own testbed on rare occasions. Several implementation weaknesses affecting carrier sense have previously been identified in standardized MAC protocols, such as 802.11's highly inefficient airtime allocation policy [Heusse03] and its limited initial number of contention slots. We are not aware of any prior work that discusses the problem referred to in Section 5 as "chain collisions".

7. Conclusion

In this thesis, we showed theoretically and experimentally that, in spite of some corner cases, carrier sense is quite effective overall and solves the spatial reuse problem in common-case, point-to-point scenarios with an efficiency approaching optimal. Fundamentally, the global reach of interference combined with the flexibility of adaptive bitrate mean carrier sense is inherently capable of providing good average performance. Good performance is achieved, and achieved under a straightforward, "average-case" threshold policy, because the smooth decay of interference with distance and the intermediate-range SNR regime targeted by data networking hardware make reasonable thresholds very robust. Through shadowing, environmental obstacles do contribute to reduced performance, but their impact is too weak to make a dramatic difference. The "hidden terminal" problem proves to be significantly less crippling than it might seem. The "exposed terminal" problem is barely a problem at all.

Our results differ from prior work in part because of the crucial attention we pay to variable bitrate, in part because we focus specifically on point-to-point, the common-case topology, in part because we consider realistic configurations and realistic degrees of concurrency rather than worst case, and in part because time has allowed radio hardware to mature, eliminating awkward quirks that plagued earlier research, such as deep fading. We show that, in short-range networks, carrier sense performs superbly. Long-range networks, owing to the increasing locality of interference, are more challenging and do include corner cases with poor fairness (on top of the unreliability already present at long range), but on average, performance is still quite good. The distinction between the two regimes, largely ignored by prior work, is a key contribution.

Where does this suggest interesting future research may lie? Multi-hop MACs remain very much an open question. Weaknesses in existing MAC protocol standards and physical carrier sense implementations also raise several questions. Improving carrier sense's corner case reliability, particularly at long range, seems a worthy goal as well. Finally, bitrate adaptation, in spite of its importance, still has room for improvement. Adaptation algorithms such as SampleRate [Bicket05], for example, will reach the optimal rate, but only as long as conditions don't change too rapidly – and they may take a while getting there. Until and unless, that is, physical layer techniques such as hybrid ARQ and rateless coding supersede algorithmic adaptation, changing the game again.

8. References

[Aguayo04] D. Aguayo, J. Bicket, S. Biswas, G. Judd, R. Morris. "Link-level Measurements from an 802.11b Mesh Network". ACM SIGCOMM 2004.

[Akaiwa97] Akaiwa, Y. *Introduction to Digital Mobile Communication*. Wiley-Interscience, 1997.

[Aoki06] T. Aoki, T. Egashira, and D. Takeda. "Preamble Structure for MIMO-OFDM Wlan Systems Based on IEEE 802.11A". IEEE 7th International Symposium on Personal, Indoor, and Mobile Radio Communications, Sept. 2006.

[Bhargavan94] V. Bharghavan, A. Demers, S. Shenker, and L. Zhang. "MACAW: A Media Access Protocol for Wireless LANs". ACM SIGCOMM 1994.

[Bicket05] J. Bicket. "Bit-rate Selection in Wireless Networks, MIT Master's Thesis, February 2005.

[Cheng06] Y. Cheng et. al. "Jigsaw: solving the puzzle of enterprise 802.11 analysis." ACM SIGCOMM 2006.

[COST231] COST Action 231. "Digital Mobile Radio Towards Future Generation Systems, Final Report". European Cooperation in the Field of Scientific and Technical Research, EUR 18957, 1999.

[Fullmer95] C. Fullmer and J.J. Garcia-Luna-Aceves. "Floor Acquisition Multiple Access (FAMA) for packet radio networks". ACM SIGCOMM 1995.

[Gallager85] R. Gallager. "A perspective on multiaccess channels". IEEE Transactions on Information Theory, vol. 31, no. 2, 1985.

[Gollakota08] S. Gollakota and D. Katabi. "ZigZag Decoding: Combating Hidden Terminals in Wireless Networks". ACM SIGCOMM 2008.

[Heusse03] M. Heusse et. al. "Performance Anomaly of 802.11b". IEEE INFOCOM 2003.

[Heiskala02] J. Heiskala and J. Terry. *OFDM Wireless LANs: A Theoretical and Practical Guide*. SAMS Publishing, 2002.

[IEEE06] IEEE Standard 802 Part 15.4. "*Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specification for Low-Rate Wireless Personal Area Networks (WPANs)*". Rev. 2006.

[IEEE07] IEEE Standard 802 Part 11. "*Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*". Rev. 2007.

[ITU-R P.1238] Recommendation ITU-R P.1238-1. "Propagation data and prediction methods for the planning of indoor radiocommunication systems and radio local area networks in the frequency range 900 MHz to 100 GHz." International Telecommunication Union, 1999.

[Jamieson05] K. Jamieson, B. Hull, A. Miu, and H. Balakrishnan. "Understanding the Real-World Performance of Carrier Sense". ACM SIGCOMM E-WIND Workshop, 2005.

[Karn90] P. Karn. "MACA - A New Channel Access Method for Packet Radio." ARRL/CRRL Amateur Radio 9th Computer Networking Conference, September 1990.

[Metcalfe76] R. Metcalfe and D. Boggs. "Ethernet: Distributed Packet Switching for Local Computer Networks". Communications of the ACM 19 (5): 395-405, July 1976.

[Patra07] R. Patra et. al. "WiLDNet: Design and Implementation of High Performance WiFi Based Long Distance Networks". 4th USENIX Symposium on Networked Systems Design & Implementation, 2007.

[Rahul08] H. Rahul, N. Kushman, D. Katabi, C. Sodini, F. Edalat. "Learning to Share: Narrowband-Friendly Wideband Wireless Networks". ACM SIGCOMM 2008.

[Ramachandran07] I. Ramachandran and S. Roy. "Clear channel assessment in energy constrained wideband wireless networks". IEEE Wireless Communications, vol. 14, no. 3. June 2007.

[Rao05] A. Rao and I. Stoica. "An Overlay MAC Layer for 802.11 networks". Proceedings of Mobisys 2005, Seattle, April 2005.

[Sridhara07] V. Sridhara and S. Bohacek. "Realistic propagation simulation of urban mesh networks." Computer Networks #51, 2007, Elsevier B.V.

[Tobagi75] F. Tobagi and L. Kleinrock. "Packet Switching in Radio Channels: Part II – The Hidden Terminal Problem in Carrier Sense Multiple-Access and the Busy-Tone Solution". IEEE Transactions on Communications, vol. 23, no. 12, 1975.

[Vaughan03] R. Vaughan and J. Andersen. *Channels, Propagation, and Antennas for Mobile Communication*. Institute of Electrical Engineers, 2003.

[Verdu98] S. Verdú. *Multiuser Detection*. Cambridge University Press, 1998.

[Vutukuru08] M. Vutukuru, K. Jamieson, and H. Balakrishnan. "Harnessing Exposed Terminals in Wireless Networks". 5th USENIX Symposium on Networked Systems Design and Implementation, April 2008.

[Yang05] X. Yang and N. Vaidya. "On Physical Carrier Sensing in Wireless Ad Hoc Networks". IEEE INFOCOM 2005.

[Zhu04] J. Zhu et. al. "Leveraging Spatial Reuse in 802.11 Mesh Networks with Enhanced Physical Carrier Sensing." Proceedings of IEEE ICC, 2004.

9. Appendix – Overview of radio propagation

Given the geometric complexity and wide diversity of environments, radio propagation is difficult to predict exactly. In known environments, rough predictions can be made using ray tracing and diffraction theory techniques (e.g. [Vaughan03], [Sridhara07]). Statistical models, however, are quite effective at predicting the overall distribution of signal strengths as a function of distance and provide valuable, high-level insights. We describe here a set of simple but well-established models, which the thesis applies to the problem of modeling carrier sense.

In empty space, radio propagation is simple: apart from a small "near field" region around the antenna (normally only a few times the size of the antenna itself), signal power decays inversely proportional to the square of distance as the wave front itself spreads out. Similarly, when a wave is confined to a plane or to a (1-dimensional) tube, the signal decays proportional to the first power of distance or not at all, respectively.

Complexity is introduced as a result of obstacles and inhomogeneities in the environment, such as walls, trees, and the surface of the earth. They may reflect, refract, diffract, scatter, and absorb the waves, and the resulting copies interfere with one another to produce the observed signal.

One of the simplest models in common use is the "two-ray" model, which takes into account the interference between waves traveling along a line-of-sight path and waves reflected from a uniform, flat earth. While this model is occasionally invoked for wireless networks, it mainly applies outdoors and in the absence of hills, buildings, and other urban obstacles. The key feature here is that, because the phase of the outgoing wave is roughly flipped when reflecting from the ground at an oblique angle, the direct and reflected waves cancel at ground level; most of the energy is in fact directed up into the air. At sufficient distance, this behavior can be approximated as a uniform decay with the *fourth* power of distance.

In more complicated environments, with many reflections and attenuating obstacles and perhaps no line-of-sight, the signal is still observed empirically decaying as a power law, but with varying exponents. Typically, exponents range from two to four [Vaughan03 p166] [ITU-R P.1238], depending on the environment. Occasionally, exponents approaching six are also found, and in long corridors sometimes they dip below two.

While power law decay provides a reasonable prediction of average signal strength versus distance, in complex environments the decay is no longer uniform. In particular, there is substantial variation from place to place, collectively known as shadowing. Empirically, shadowing is known to follow a lognormal distribution. A simple intuition is as follows: For any given region of space, a large number of obstacles and reflections may or may not contribute to shadowing, and each possible obstacle can be considered a binary random variable. The contribution of each obstacle is a small multiplicative attenuation factor, and the net shadowing loss is the product of all these attenuations. Or, on a logarithmic scale, it is their sum. By the central limit theorem, this sum of small random variables tends to a Gaussian. In typical environments, the standard deviation is usually found around 4dB to 12dB [Akaiwa97] [ITU-R P.1238] [COST231 §4.7.6].

Note that the variation described by the shadowing model goes both positive and negative. In part this reflects the fact that "shadowing" can provide a positive contribution from reflections,

but more importantly, it's because the obstacles that cause shadowing also contribute to the overall, empirical path loss exponent, and so positive deviations can represent merely a localized reduction of loss. For example, in front of a wall, there might be a positive deviation, while behind it, a negative one. Path loss captures only the large-scale effect of the presence of walls.

The same environmental complexity that leads to the shadowing effect at large spatial scales also leads to an important small-scale and frequency domain effect known as multipath fading. The assortment of reflected waves creates a complicated, frequency-dependent interference pattern that affects amplitude both as a function of position and frequency. The amplitude distribution is known empirically to follow a Rician distribution in the presence of a line of sight to the transmitter and a Rayleigh distribution in the absence of a line of sight. The intuition is as follows: Each of the reflected components arrives with a random phase and amplitude. When visualized in the complex plane, each component is a short vector pointing in a random direction, and the total signal amplitude is the sum of these vectors. By the central limit theorem, the total signal vector then follows a bivariate Gaussian, centered at the origin in the absence of a line of sight, or offset somewhere else in the presence of one. The amplitude of the signal then follows the distribution of the distance of the signal vector from the origin. For a centered bivariate Gaussian, this is a Rayleigh distribution, and for a non-centered bivariate Gaussian, it is a Rician.

Rayleigh fading, in particular, has several important effects. First, for channels whose bandwidth is wide enough to resolve the varying delays of different echoes, the reverberation causes adjacent transmitted symbols to overlap and interfere; this is known as inter-symbol interference (ISI). ISI is in large part an effect of phase distortion and is reasonably straightforward to reverse (with adaptive equalization, for example) or engineer around (with OFDM-style modulation), although a delay spread greater than anticipated by the design can still cause catastrophic link degradation (e.g. [Aguayo04]) which cannot be resolved by increasing transmit power.

Second, besides the frequency-dependent phase distortion, there is also frequency-dependent amplitude distortion, known as "frequency-selective fading". This effect can never be entirely corrected; while the deeply faded frequency bands can be re-amplified, doing so also amplifies the background noise. Thus, radios must find a way to tolerate frequency-selective fading, typically by adding some sort of redundancy (often called "diversity").

Narrowband radios have few options: They can either hop around from frequency to frequency (as in FHSS), use multiple antennas spaced apart by a significant fraction of the wavelength (to take advantage of the spatial dependence of the fading pattern), or simply accept an increased rate of random outages and capricious connectivity due to wide variance in RSSI.

Wideband radios, on the other hand, have inherent diversity in their use of a broad swath of frequencies. OFDM radios can code across different sub-carriers (as in 802.11a/g) or use adaptive modulations on a per-sub-carrier basis. DSSS radios can use the RAKE receiver, a form of adaptive equalizer, to coherently combine energy from different echoes (as in CDMA cellular networks and most 802.11b receivers). Such wideband techniques, sometimes in combination with multiple antennas, help to dramatically reduce the unreliability inherent in the fading channel. From a capacity perspective, it reduces to the equivalent of a few dB variation, at which point that we can largely ignore it compared to the effects of shadowing.

Figure 14 below illustrates an example application of this model to an indoor 802.11 testbed, along with measured RSSI figures.[20] Each data point represents the receiver SNR for one (detectable) sender-receiver pair in the testbed. The plotted curves show the mean and +/- 1 standard deviation bounds for the maximum-likelihood fit of a model combining power law path loss and lognormal shadowing (and accounting for the invisibility of sub-threshold links). Because the measurements are relatively wideband, the residual variation due to Rayleigh fading is small enough to be ignored compared to shadowing. The model shows a rather good fit for the experimental data in the vertical direction, insofar as quality of fit is well-defined given the wide shadowing variance. (The model does not attempt to predict the horizontal distribution – this is in large part a function of the layout of the testbed nodes.) The maximum-likelihood parameters are $\alpha = 3.6$, $\sigma = 10.4$dB, and $RSSI_0(R=20) = 46$dB.

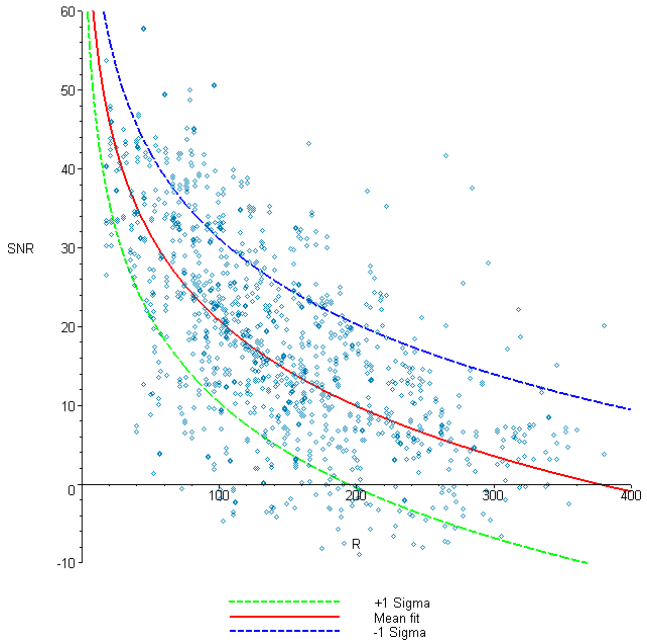

Figure 14 – Plot of measured wideband signal strengths at 2.4GHz for all pairs in the testbed, along with a maximum likelihood fit of the path loss / shadowing model, showing +/- 1 standard deviation bounds. Vertical axis is in dB, while horizontal axis is in somewhat arbitrary map coordinate units. $\alpha = 3.6$, $\sigma = 10.4$.

[20] SNR figures are calculated form RSSI values measured with 1Mbps broadcast packets, with attempted corrections for variation in hardware noise floor using a least squares fit over the constraint of symmetric channels given the assumption of uniform transmit powers (no hardware variation in transmitter strength). Uncorrected data is largely similar. Note that these data points cannot be directly compared with the experiments of section 4, because those were conducted at 5GHz, where, unfortunately, no 1Mbps rate is available for sensitive RSSI probes.