IEEE P802.11
Wireless Access Methods and Physical Layer Specifications

ELABORATE CLEAR-CHANNEL ASSESSMENT
FOR INDOOR COMMUNICATION SYSTEMS
OPERATING IN UNCONTROLLED UHF & MICROWAVE BANDS

Date: July 8, 1994

Author: Lawrence H. Zuckerman
290 Gaffney Hill Road Easton PA 18042
Mail: PO Box 828 Easton PA 18044 USA
(610) 252-1587

ABSTRACT

An overall strategy--both technical and legal--for developing the IEEE P802.11 Clear-Channel Assessment ("CCA") standard is recommended to the Joint MAC/PHY Standards Committee. It is suggested that, in order to obtain the maximum network throughput, the decision of when to transmit and when to hold transmission in abeyance should be based upon many factors and a convoluted process, and is deserving of more resources than is generally supposed. It is further suggested that the specification of this process should be decided upon only after much testing and analysis. The issue of whether or not CCA should be managed by the MAC Layer is also discussed. A novel approach is advanced, which purports to eliminate what is commonly considered to be a fundamental CCA limitation. It is demonstrated how a sophisticated, computer-intelligent approach such as this one could materially increase network throughput.

A. Introduction

Perhaps the most critical (and most difficult to design) function of the Wireless LAN types soon to become part of a P802.11 Standard ("Standard") is Clear-Channel Assessment ("CCA"). The Author's proposed goal of optimal CCA is to determine transmission start times which allow, on the average, largest possible overall network throughput and minimal interference to all intentional radio activity. Even under the best circumstances, dealing with the vagaries of indoor propagation, multiple networks and multiple services is already difficult. In addition, the dual (self imposed) requirements of "Carrier" Sense (rather than Polled) Multiple Access and physical peer-to-peer communication impose additional severe burdens on CCA. Only the most carefully optimized criteria will permit the best achievable network throughput (under the circumstances). It will be necessary to employ every possible clever technique available. The intent of this Paper is to suggest some of the approaches that might be considered by those who will be preparing additional CCA proposals for presentation in July. (At that meeting, we must make a final decision.)

Summary:
Before beginning the discussions, a number of terms are defined. This precaution is particularly important for the present topic, in which so many concepts that were initially invented for wired communications have been inappropriately applied to the wireless media (and the MAC Engineers will be delighted). Some of the definitions were coined by the Author for the present submission; other definitions appear elsewhere. Next, some of the submissions and discussions on CCA that have already taken place are reviewed and subjected to rather detailed comment. These comments strongly indicate that multifaceted CCA is needed for maximum network throughput. One comment is quite provocative; it contradicts what is generally supposed to be a fundamental limitation of the CCA process. All of this discussion sets the stage for Part D wherein the method for providing "true" CCA is explained, and then tested for usefulness by a calculation. It is valuable to have creative ideas, but that is insufficient in a practical context. Therefore, Part E lists the Issues which must be settled (especially in light of the new ideas), in order to complete a competent CCA Standard. The Conclusions merely state that if the job is not done correctly now it will be done later at far greater cost.

B. Definitions

In order to minimize confusion and the consequent unnecessary discussions, it is more than just useful to employ the following definitions:

1. The subject "Band" is whichever one needs CCA, but especially the 2.4 GHz ISM band.
2. A "Radiator" is any device which emits RF energy in the Band.
3. A "Station" is any intentional Radiator.
4. A "Compliant Node" is Any equipment manufactured in accordance with a P802.11 Standard and operating in the Band.
5. The "Subject" Node is the one performing CCA, because it has traffic for the "Recipient".
6. Any Compliant Node which is distinguishable as such by the Subject node shall be referred to as a "Compatible" Node. All other Nodes are termed "Incompatible". (Under this definition, a direct sequence Node is Incompatible with a Subject frequency hopping Node, if the latter cannot distinguish between signals from the former and an increase in the noise floor.)
7. Any Compliant Node which is in the same network as the Subject Node is called "Native". All other Nodes are "Foreign".
8. The "Channel" is whatever segment of the Band to which the Subject Node's receiver is currently sensitive. This Band segment might be small for a narrow band or frequency hop Station, or large for a direct sequence Station.
9. An "Unoccupied" Channel refers to no detectable signal whatever by the Subject.
10. A Channel is defined herein as "Occupied" or "Active" if the Subject RSSI registers any reading above that which would be obtained with a matched load substituted for the antenna.
11. Any detected RF Power is referred to as a "Signal".
12. Two or more Signals detected simultaneously is a "Collision".
13. A "Transmission" is any signal from an intentional radiator.
14. From any Recipient's point of view, a "Readable" Transmission is one that can be

decoded with little enough BER to be useful, and the "Desired" Signal is the first Readable Transmission addressed to it and still in progress.

15. A barely Readable Signal is called a "Threshold" Signal.
16. A Transmission from a Compatible node is called a "Packet" or "Fragment".
17. A Transmission is said to be "Existing" if it is being detected by the Subject during a period when a Packet or Fragment is ready. The Node responsible for this Transmission is called the "Incumbent".
18. A channel is deemed "Clear" whenever a Transmission would begin forthwith (for whatever reason) if a Packet were ready; as in "Clear to Send".
19. Whenever a Transmission is postponed during a period when a Packet or Fragment is ready, the channel is deemed "Busy", and the decision to postpone is referred to as a "Deferral".
20. If a Deferral is made, or extended, but in reality; a) the Transmission would have been received without error by the node for which it was intended, b) the acknowledgment could have been received, and c) neither the transmission nor its acknowledgment ("ACK") would have "Ruined" (caused an error in an otherwise error free) Native Transmission; then the entire Deferral or that portion of it as the case may be is called a "False Deferral".
21. If a Transmission is made while the Channel is Occupied and it Ruins a Native Transmission which would also have completed ACK, it is termed "Detrimental", whether or not it and its ACK were completed.
22. Supplementary information (to be explained in this Paper) sent during each Transmission--*after* the bit patterns which are necessary to ascertain the correct receive antenna, acquire bit synchronization, and define the start of data fields--but *before* start of the payload's header, will be termed "Pre-data".

Other distinctions, such as the various Detected Signal strengths of an Occupied channel, will be made in connection with the Clear channel decision. Notice that there is no distinction between signals Per-Say and noise, inasmuch as a low cost FH LAN receiver which recognizes a DS signal as other than a noise floor increase is probably not practical at the present state of technology.

Other definitions will be made in subsequent sections where they can be more clearly explained by the context. In all instances, terms previously defined are capitalized. In some instances, these terms are followed by the definition Number in brackets.

C. Comments on Existing Recommendations

In writing the present submission, the Author (who has designed, developed, and directed the manufacture of wireless network equipment which involved critical CCA considerations and attended all P802.11 sessions during the past year) has the benefit of CCA discussions at the present and past two sessions and of the CCA submissions by Kawaguchi/SYMBOL (94/70) and McDonald/MOTOROLA (94/79, 94/110). It is recommended that the Reader review or study these submissions before continuing his perusal of this one, and that he carry along copies of all CCA materials to the July meeting. In this section, the Author's *comments*, as opposed to his summaries, are italicized. This practice not only eliminates confusion but allows review and comments to be efficiently intermingled.

Analysis of IEEE P802.11-94/79

McDonald's first submission is an excellent tool for familiarization with wireless CCA. He first explains the importance of CCA and reviews two of the basic differences affecting the CCA problem in wired and wireless systems. His overall treatment is to rate four CCA approaches ("packet" detect, RF power detect, symbol rate detect, and a combination of the last two) in accordance with how well they relate to five important aspects of radio packet networks (propensity for False Deferral, hidden node syndrome, communications delay, compatibility with Native Nodes of different speeds, and etiquette with respect to Foreign Nodes).

The foundation for McDonald's reasoning comes from his three diagrams, which brief the Reader on with what the CCA process must contend. The first diagram shows the Subject Node, the Recipient Node, another Native Node whose Transmission is Occupying the Channel, and his Recipient Node. *Inasmuch as the stated basic goal of CCA is to promote maximum network throughput, all (equal length) Fragments have the same importance; so the decision not to Defer must be based upon the total probability that*

- *The Subject's Fragment will be received correctly;*
- *The Subject's ACK will be received correctly;*
- *Neither of the above Transmissions will Ruin the Existing Transmission (or its ACK).*

The key factor, which governs the stated limitations of the CCA methods is that the diagrams are hypothetical; nobody knows the true locations of these four Nodes relative to one another.

In his next section, the four CCA forms are described in detail.

RF Power detect is assigned two facets: A threshold which has been taken by others as a reasonable compromise (12 dB above the 10^{-5} BER level) is stated without specific endorsement. An approximate measurement period (50 µsec) is termed "short". Little detail is given; as Power Detect is eliminated later in the submission for superseding reasons.

Considering the position uncertainty of all Native Nodes and the number of different scenarios that would lead to a Detrimental Transmission, under these conditions the threshold should be much lower indeed. (This same comment applies to his "hybrid" form.) If the measurement period refers to how long it would take for RF appearing at the antenna to register as a CCA output, 50 µsec is an eternity for network designers; as during this entire period, the Subject Node could begin a Transmission on an Occupied Channel.

Bit or symbol rate detect at least eliminates all Incompatible Nodes from consideration. *Like all clock recovery functions,* it is considerably more sensitive than data recovery.

Packet detect is the only CCA form which distinguishes Native Nodes from all other Radiators, thus eliminating them from consideration. In order to distinguish Native Nodes from Foreign Compatible Nodes, it is necessary to decode the header portion of the transmission.

Even if the Subject Node monitors all the time it is not transmitting, the header may be obscured; so in order to minimize Detrimental Transmissions, packet detect should be combined with one or more of the other CCA forms.

The section detailing performance comparisons and the remainder of this submission are a bit convoluted; therefore, only the material needing commentary will be reviewed. The

second diagram is a special case of the first one, illustrating the hypothetical positioning of Compatible Nodes such that if the Subject Node (T) did not Defer to an Existing Transmission from a Native Node (I), the T's Transmission would be successful. One intention here is to show that a False Deferral is the result; as T has no way of knowing about this otherwise fortunate positioning.
Absent from this diagram is I's Recipient Node, who must typically be in his opposite direction than T in order to make the False Deferral complete.

It is stated that Deferring to microwave ovens and Foreign Nodes (even Compatible Nodes) has little utility. The rationale for disregarding even Foreign Compatible Nodes is that they are generally far enough away from either the Subject Node to allow the isolation of spread spectrum to minimize the number of Ruined Transmissions.
If the microwave oven or any other Radiator is strong enough at the Subject's location to prevent the latter from receiving an ACK, (unless the Recipient is the only other Native Node or an emergency exists) he should <u>never</u> transmit. If he does, he will re-transmit over and over again, running down his battery; and each transmission will carry the potential to Ruin others. In many situations, networks using Compatible Nodes will be operating in close proximity. In those instances, Ruined Transmissions will be common. The protocol will revert to Aloha, seriously limiting throughput unless the networks are lightly loaded. (McDonald's protocol <u>is always</u> Aloha with respect to Foreign Nodes.)

The last diagram shows the famous Hidden Node Syndrome. *As expected,* symbol (clock) detect is shown as more sensitive than the packet (data recovery) method. Based (*only*) upon the threshold setting discussed earlier, RF Power detect is deemed least sensitive, and therefore least able to cope with Hidden Node.

McDonald reasons that in this uncontrolled Band, protocols and modulation systems are so diverse that the FCC provided for built-in etiquette by the spread spectrum requirement. Therefore, Deferrals to Foreign Nodes are False; so any CCA approach which Defers to Foreign Nodes should be eliminated from further consideration. Thus, RF power detect alone is deemed not useful; as it cannot single out Native Nodes.

Handling Native Nodes with different speeds turns out to be a possible problem. RF power detect has been eliminated. Clock recoveries designed for the anticipated GFSK will not respond to all the modulation types being discussed for the higher speed FH PHY. The higher speed FH PHY will use a fully Compatible preamble; so the Packet Detect method works to Defer *if and only* if the Subject decodes it *and it knows (from information in the preamble) the length of the High Speed transmission. Fortunately, this problem was solved as far as Symbol Rate detect is concerned; as the High Speed FH PHY will apparently use a fully compatible 4FSK.*

Overall, Packet Detect will yield the best Network throughput. RF Power Detect is far too susceptible to False Deferral and (*as the threshold is defined here*) does not handle Hidden Node as well as the other two basic types. Symbol Rate cannot distinguish between Native and Foreign Nodes, which impacts on False Deferral.
<u>Given the tools stated as available in this Submission</u>, Packet Detect should be the primary CCA method. However, considerable throughput improvement will be obtained if all three basic methods are intelligently used.

1) Bit / Symbol Rate is needed to assist in the identification of Packet portions other than the preamble, especially higher speed FH. 2) Symbol Rate, combined with RF Power Detect, is also needed to identify proximal Compliant networks. 3) Symbol Rate may even allow Narrow Band and Frequency Hopping receivers to detect Direct Sequence.
1) RF Power Detect is needed to determine that the Signal is too strong to receive an ACK or determine the status of the Subject's network. 2) As explained previously, it is important to avoid beginning a transmission when the Channel is actually occupied. If the start of a Signal is in transit (propagation delay), there is no way to recognize its existence, and a Collision will occur. Inasmuch as it is the quickest indicator of Activity, an RF Power Detect should result in a temporary Deferral until the other method indicators have time to register. 3) It is also needed to analyze time patterns of the Signals to determine the best-bet Transmission Lengths and start times. 4) Finally, it can also be used with variable threshold, in conjunction with the types of Signals Occupying the Channel or Channels, how long the Transmission has already been Deferred, and needs of the upper OSI levels.

The fundamental problem of CCA is that the Subject Node can be cognizant of the interference conditions only at his location, yet an accurate CCA can be performed only if one knows the conditions at his Recipient's location.
To perform an accurate CCA, the Subject needs to know the conditions at his location, his Recipient's location, the Occupying Node's location, and his Recipient's location; and using the proper techniques, all of this can actually be accomplished to a close enough approximation to materially improve throughput of the Native network and all surrounding Compatible networks.

COMMENT
Unfortunately, the Author was unable to devote the time to thoroughly analyze the following Submissions:

IEEE P802.11-94/70
Kawaguchi at least recommended CCA to be based upon more than one indicator.

IEEE P802.11-94/110
McDonald's second contribution shows the benefit of discussions at the March meeting, additional considerations, and receipt of definitive information about microwave oven radiation.

IEEE P802.11-92/125
Larry Van Der Jagt submitted this short paper, which shows how a well thought out MAC/PHY interface can facilitate flexible and elaborate WLAN node functions.

D. New Recommendations: Elaborate CCA

New Definition: TDA instead of CCA

Before beginning the Part in earnest, it is useful to define a new term. In radio work on uncontrolled bands there is, in general, no clear channel; so Clear-Channel Assessment is really a misnomer. The process is to use a complicated formula based upon data on the character of the received Occupying Signal and information from upper OSI layers to determine whether or not to Defer. Therefore, the new name for this process is TRANSMISSION DEFERRAL ASSESSMENT or "TDA".

Introduction

In Part C, the Author states that considerable further improvement in LAN operation would result if TDA were carried out by the Subject Node with some knowledge of conditions at relevant other locations. It was also asserted that there is actually a method to obtain this information. Before disclosing this method, it would be helpful to discuss more graphically what this information would mean to the LAN designer. Refer again to the three Node Position diagrams in Jim McDonald's 94/77 Submission. Suppose these "maps", instead of representing possible arrangements of Nodes to illustrate TDA limitations, represent any actual present disposition of Nodes, of which all Compatible Nodes within communications range are fully "aware". The Subject Node could use this (and other) information to determine with relatively high accuracy whether he can make a Transmission to his Recipient and receive an ACK with an acceptably low probability of Ruined Fragments, even though the Channel is Occupied by another Node.

The germination for this idea came from a discussion at one of the FH PHY working sessions two weeks ago in Oshawa, Ontario. Peter Chadwick (PLESSEY) pointed out that knowledge of transmit power in addition to received signal strength allows one to estimate the distance to Compatible nodes. I tried to underscore the importance of this idea, but nobody seemed to be impressed.

The "True" TDA Method

Under this plan, it would be necessary to display Effective Radiated Power ("ERP") in every Fragment preamble, with a resolution of 1 dB. If the overall range is from 1 mW up to 1.6 W, only 5 bits are required. Variable transmit ERP on a per-packet basis would provide additional throughput improvement. 3 dB steps would be ideal, but any ERP variation would be helpful. By keeping (initiating and updating) a ("Chart") of transmitted and received ERP for every readable Node, each Node would have rough distance data (i.e. could draw concentric circles representing the distances). More precisely and importantly, each Node would "know" the path loss between himself and every other Node.

The other part of the plan is that, in some prescribed manner, each Node periodically transmits its Chart; so that every Node ends up with all the Charts. The system determining when a Node includes his Chart (or even on occasion another Node's Chart, to eliminate the Hidden Node effect) as a field, is governed by many factors, such as network loading and perceived changes to his Chart. During periods of light loading, the Charts can be broadcast via special network maintenance Transmissions; so they are ready for efficient operation when payload Packets/Fragments are ready.

To quickly understand how this can work, it is only necessary to realize that a map showing the relative positions of any four non co-linear points in a plain can be drawn using only the scalar distances between each of the points. The Subject Node does not need (and would not want) to draw a map; it has the exact information it needs to perform TDA:
1. Incumbent Node's ID, ERP, and Recipient's ID, from Incumbent Node's preamble;
2. Path loss between the Incumbent Node and Recipient;
3. Path loss between himself and his own Recipient;
4. Path loss between himself and the Incumbent Node's Recipient;
5. Path loss between the Incumbent Node and Subject's Recipient.

All Subject has to do to begin a Transmission is calculate the lowest ERP which makes him 15 dB or more stronger than the Incumbent Node at his own Recipient's location (and strong enough to be decoded by Same) and 15 dB or more weaker than the Occupying Node at the latter's Recipient. If no such ERP is available to Subject, he Defers!

The method works just as well for 3 dimensional geometry.

The method will increase network throughput, because it will decrease the probabilities of both False Deferrals and Ruined Transmissions. Owing to the available information, there is a greater probability of finding a high enough ERP; as every Transmission, including the ones initiated with no Activity, can be made using minimal ERP.

Estimation of Network Efficiency Improvement

It is possible to calculate throughput improvement by making several conceptual simplifications. It is important to bear in mind that the actual system, which measures path loss under field conditions, will do a much better job of determining signal levels than these calculations. The calculation shown here applies to strict "Packet Detect" as a baseline. The specific question to be answered is: "By deferring to all Packet Detect Incumbents, what percentage of these Deferrals are False?" The assumptions are:
1. In order for the Desired Transmission not to be Ruined, it must be at least 15 dB stronger within the information bandwidth than the total of all other signals and noise.
2. The maximum Effective Radiated Power (ERP) to be used is +20 dBm. (This power level is consistent with the need for survival of PCMCIA power circuitry in notebook sized computers.)
3. The Subject Node is in a "Sea" of Native Nodes having a radius at least twice as far as his Transmissions are Readable.
4. A Node has the same probability to be in any position.
5. The path loss formula is the same one used by Allen in IEEE P802.11-93/105, quoting Tuch in IEEE P802.11-91/69:

FORMULA 1

where: G is the path loss in dB, d is the distance in meters from the RF source, and λ is its wavelength. This formula was derived using the assumption that for indoor propagation, the attenuation for the first 8.5 meters from the source is similar to free space conditions (square law of power flow, represented by the second term), and that for all distances greater than 8.5 meters there is on average a 3.6 law of power flow (represented by the first term). The last term is the correction required which takes into account that the smaller the wavelength, the smaller the solid angle can be captured by an antenna having a reference gain value.

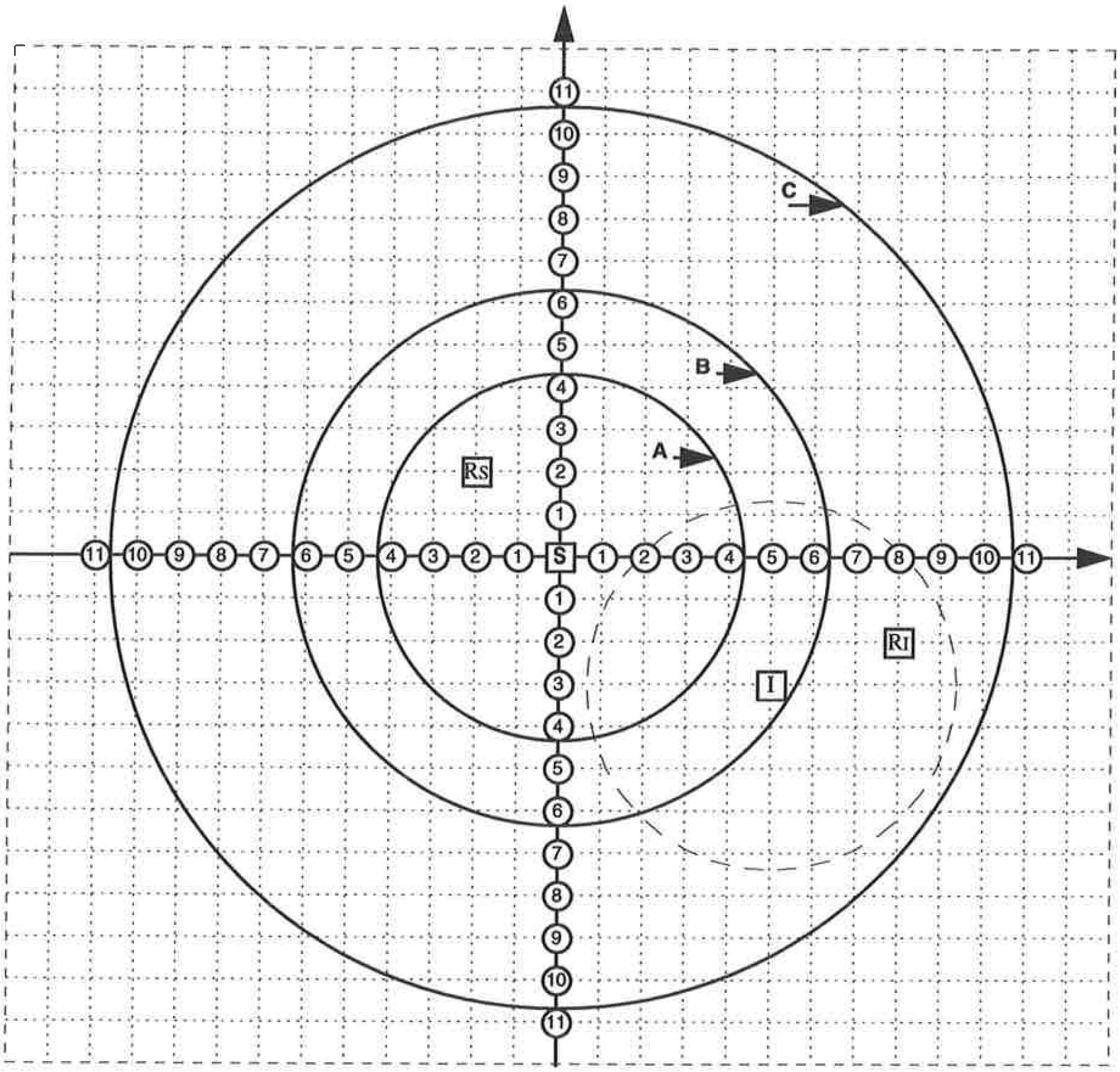
As explained above, given that the facilities and procedures make the following statements true:

1. Each node has knowledge of the path loss between every node and every other node;
2. Every Fragment is transmitted with ERP which delivers a signal to the Recipient which is 15 dB stronger than Interference + Noise within the information bandwidth;
3. Every Transmission contains Pre-data [22] that includes the current ERP value;
4. At the time a Fragment is ready for transmission, if there is an Incumbent (and there will be more than half the time while the network is heavily loaded), the Subject will initiate its transmission if and only if there exists an ERP between -10 dBm and +20 dBm such that (both):

****His signal will reach his own Recipient Node at least 15 dB stronger than that of the Incumbent;*
****His signal will reach the Incumbent's Recipient Node at least 15 dB weaker than that of the Incumbent.*

The objective of this TDA Calculation is to estimate the percentage of cases that the Subject's Transmission could actually be made, Ruining neither his nor the Incumbent's Transmission. In other words, using Packet Detect and equal ERP for all Nodes, what percentage of Deferrals are False? Inasmuch as a Node has equal probability to be anywhere, the analysis can be performed by a method of areas. Such analysis automatically results when considering a sea of Nodes having uniform density.

(Figure 1 is on the next page)



IEEE P802.11-94/132 Fig 1: "SEA OF NODES" L. H. ZUCKERMAN

Figure 1 shows the Subject Node "S" in the center of several imaginary concentric circles and surrounded by numerous nodes, one at the intersection of each dotted line. As a particular example, three other nodes are marked on the Figure:

- "R_S" is the intended Recipient of a Fragment from "S", which is ready to be transmitted at the present time.
- "I" is the Incumbent Node, already transmitting a Fragment at this time.
- "R_I" is the Incumbent Node's Recipient, now receiving its Fragment.

The innermost Circle (**A**) has a radius equal to the path loss distance at which (for a given specified ERP and receiver sensitivity) there is a Threshold signal. Therefore, this circle contains all the Nodes with which the Subject can communicate directly, assuming there is no interference. For the case of GFSK radio modems operating at 1 Mb/sec, we shall assume that a signal to noise (within the information bandwidth) ratio of 15 dB will yield a

Bit Error Rate ("<u>BER</u>") of about 2 x 10^{-6}, which is chosen so that the error rate (before any re-transmissions) of the largest Packet or Fragment (4800 bits) will be 10^{-2}. (This is also known as the Block Error Rate.) Retransmitting only 1 long-and-distant Fragment out of 100 probably represents an overall network retransmission rate of 10^{-4}, assuming on average one Fragment out of ten is long and one fragment out of ten is sent to a Node at this "Threshold" distance. This result is splendid, (but in practice, the network performance is much poorer owing to multipath fading and interference).

To determine the actual radius of Circle **A**, we first compute the effective noise level within the information bandwidth (which we shall take to be 750 KHz). This bandwidth is 59 dB above -174 dBm/Hz, or -115 dBm. Assuming a system noise figure of 6 dB, the effective noise level is -109 dBm. According to the above assumption about equipment sensitivity, the signal level at the receive RF input must be 15 dB above this, or -94 dBm. Assuming unity receive antenna gain and using the ERP figure of +20 dBm, the path loss is 114 dB. For slow frequency hopping, the effects of (Rayleigh) fading statistics must be factored in. Assuming two-antenna diversity, a 12 dB margin seems sufficient; so the base path loss can be only 102 dB. Using the above path loss formula, this works out to a radius of 135 meters, or 443 feet (shown on the Figure). *"R_S" must be located within this circle.*

The next larger Circle, **B**, encloses all possible locations for the Incumbent Node, "I". The Subject Node is able to identify nodes further away than it can effectively communicate with them; as the Pre-data [22] portion of each Fragment (containing, among other things, I and R_I node identification numbers and ERP) is much shorter (say 64 bits) than a long transmission. Therefore, the BER can be much higher. For a BER of 10^{-2}, one half (on the average) of any randomly chosen sets of 64 bits will be error free. Using the BER vs. S/N family of curves, where N has Gaussian statistics, S/N drops by 5.4 dB as BER changes from 2 x 10^{-6} to 10^{-2}. Adding 5.4 dB to the above path loss formula increases the radius to 191 meters or 625 feet.

Circle "**C**" encloses all possible locations of R_I. All four relevant nodes are located within this circle. It reflects the geometry that any I at the edge of **B** can communicate with an R_I as far away from **S** as <u>**B** radius</u> + <u>**A** radius</u>, or 1068 ft.

The off-center dashed-line circle defines the region within which R_I can be located for the particular location of I shown at its center.

In order to actually carry out the TDA Calculation, it is useful to diagram the path loss between any given Node and all the other Nodes contained within the Circles shown on Figure 1. Figure 2 shows this picture for conceptual clarity, and allows one to read directly the path loss between the Subject and any other Node.

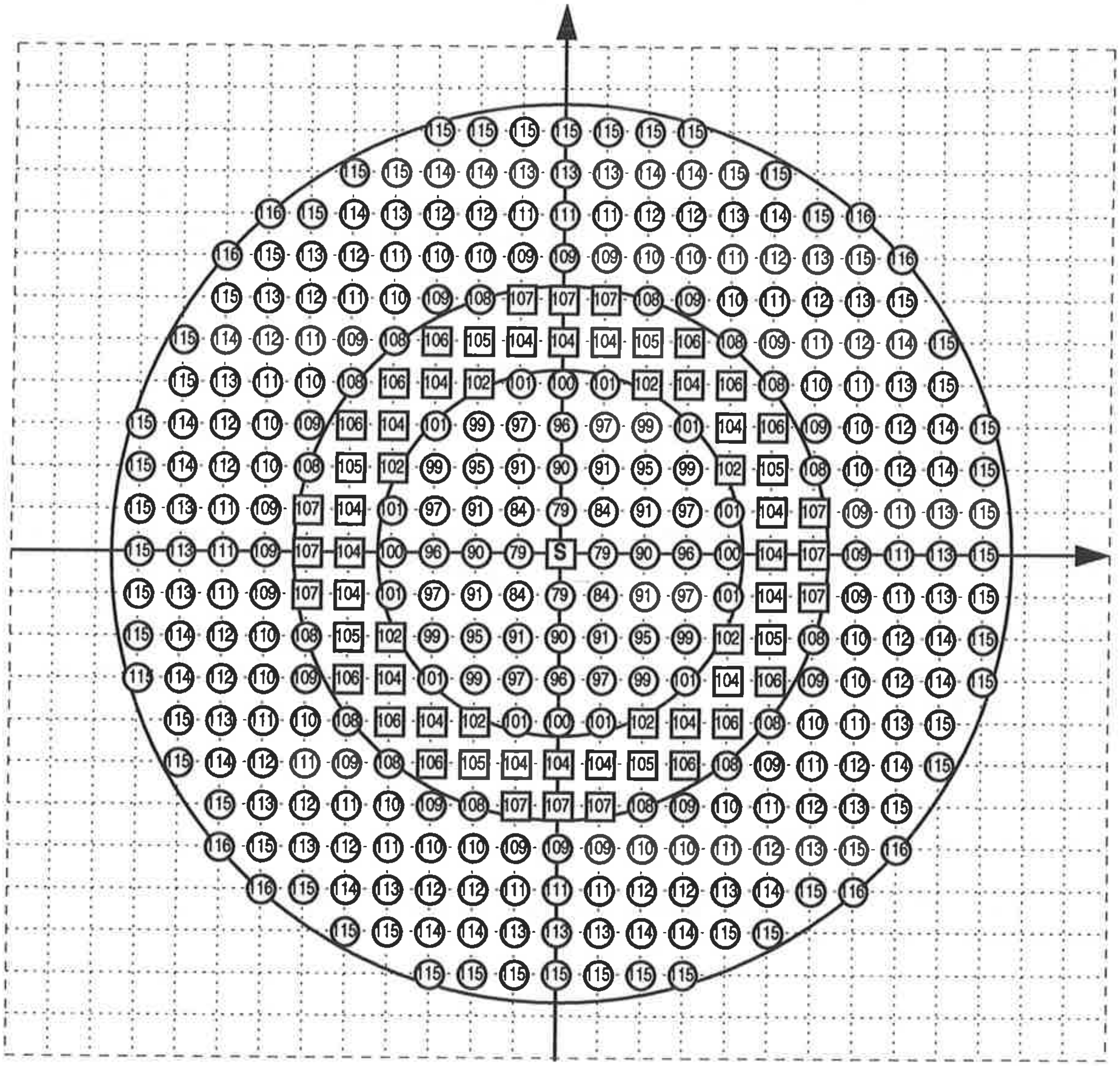


Figure 2: "PATH LOSS"

This Figure also shows which Nodes are located within Circles **A**, **B** (the ones with the path loss numbers shown in squares plus the ones within **A**), and **C** (all Nodes), but the following table is more useful for calculation and is far more compact:

• • •	0	1	2	3	4	5	6	7	8	9	10
0	0	78.7	89.6	95.9	100.4	103.9	106.8	109.2	111.3	113.1	114.7
1		84.2	91	97	101	104	107	109	111.4	113.2	114.8
2			95.0	99	102	105	108	110	111.7	113.5	115.1
3				101.3	104	106	109	110	112.3	113.9	115.4
4					105.8	108	110	111	113.0	114.5	
5						109.3	111	112	113.8	115.2	
6							112.2	113	114.7		
7								114.6	115.7		
8											

TABLE 1

This table allows one to quickly determine the path loss between any intersection on Figure 1 and any other intersection closer than 1000 feet. To find the path loss between any two nodes on Figure 1, count the number of intersections along either axis, then the other axis. The cell in this table intersecting the boldface row and column numbers equal to the number of intersections contains the path loss. For instance, to find the path loss between the Incumbent Node, I, and the Subject's Recipient Node, we count 7 intersections on one axis and 5 on the other, and read 112 dB on Table 1.

Calculation Procedure

The last assumption needed for this analysis is that every Node has an equal probability to communicate with any other Node within its range, and with no bias in Transmission length. In order to estimate the improvement to be gained in network throughput by this particular Elaborate TDA method; as opposed to the simple Packet Detect method, it is necessary to compare the number of combinations for which I, R_I and R_S can be situated that allow S to transmit, as opposed to the number of combinations for which S must defer.

The first of two parts of this calculation is to derive the formula which relates the four relevant path loss numbers obtained for each combination of Node locations to the Transmit/Defer decision (which includes the permitted S Node power range). As mentioned above, the path loss numbers are:

1. G_{II}: between the Incumbent and its Recipient;
2. G_{IS}: between the Incumbent and the Subject's (intended) Recipient;
3. G_{SI}: between the Subject and Incumbent's Recipient;
4. G_{SS}: between the Subject and its (intended) Recipient.

As also mentioned, the Transmit/Defer decision relates to:

1. Maximum allowable ERP of +20 dBm;
2. Minimum and target received Signal to Noise Ratio (15 dB);
3. Minimum allowable Signal to Interference Ratio (15 dB).

In order to derive this formula and conceptualize the problem, let us perform the TDA for the particular combination of Node locations shown in Figure 1.

1. The three coordinates (for I, R_I, R_S) are 5,-3; 8,-2; & -2,2.
2. The four path losses (G_{II} through G_{SS}) are: 97, 112, 112, 95 dB respectively.
3. The signal level at R_I from I (defined as P_{II}) is always -82 dBm, which is 15 dB above -109 dBm noise, plus another 12 dB for Rayleigh fading statistics with 2-antenna diversity.
4. We then compute the ERP of I (defined as P_I):
   $P_I = P_{II} + G_{II} = -82 + 97 = +15$ dBm. $P_I </= +20$ dBm.
5. Note: Step 3, in facilitating the present calculation, assumes I initiated its Transmission during Clear channel conditions, which is obviously not always the case. Thus, in practice, -82 dBm is only a minimum P_{II}. The field equipment would do the Step 4 computation in reverse; it would receive P_I as part of I's Pre-data and determine P_{II} using the value of G_{II} currently in its Chart.
6. Next, we compute the maximum ERP allowable for S ($maxP_S$) in two steps. Start by subtracting 15 dB from P_{II} to obtain P_{SI}, which is the maximum allowable power from S incident at R_I:
   $P_{II} - 15 = maxP_{SI}$; or -82 dBm - 15 dB = -97 dBm.
7. Then, we add the path loss between R_I and S:
   $maxP_S = maxP_{SI} + G_{SI}$ = -97dBm + 112 dB = +15 dBm. $P_S </= +20$ dBm.
8. Next, we find out whether this P_S is enough to cover the path loss to S's intended Recipient:
   $P_{SS} = P_S - G_{SS}$ = +15 dBm - 95 dB = -80 dBm,
   which is >/= -82 dBm; thus passing this test.
9. Finally, we determine whether the power incident at S's intended Recipient from S is at least 15 dB stronger than that from I in two steps:
   $P_{IS} = P_I - G_{IS}$ = +15 dBm - 112 dB = -97 dBm.
10. $P_{SS} - P_{IS} - 15$ = -80 dBm-(-97 dBm) - 15 dB = +2 dB; which is >/= 0 dB, thus passing the final test.
11. THEREFORE, S should not defer; it should transmit immediately, with ERP of +13 to +15 dBm.

Using the various equations and inequalities generated in the above analysis, the following simple formulas were derived for testing the path loss values resulting from every possible combination of I, R_I and R_S Node placement on Figure 1. First, the basic (interrelated) constraints are repeated:

1. P_I & P_S are both less than or equal to +20 dBm;
2. G_{II} & G_{SS} are both less than or equal to 102 dB;
3. P_I equals G_{II} plus (-82 dBm).

Next, the test formulas are given. If the following statements are true, S can transmit, thus avoiding a False Deferral:

1. $P_S >/= G_{II} - G_{IS} + G_{SS} - 67$ (dBm)
2. $P_S </= G_{SI} - 97$ (dBm)
3. $G_{SI} - G_{SS} >/= 15$ (dB)

Repeating the above example using these inequalities:
P_S >/= 97 - 112 + 95 - 67---+13 dBm; P_S </= 112 - 97---+ 15 dBm; 112 - 95 = 17 dB.
The Subject Node is permitted to transmit using any power level between +13 and +15 dBm.
As another example, if, for instance, R_S is moved one intersection to the left in Figure 1, to (-3,2), the path loss numbers become 97, 114, 112, & 99. The formulas then report:

P_S >/= 97 - 114 + 95 - 67---+11 dBm; P_S </= 112 - 97---+15 dBm; 112 - 99 = 11 dB.
The Subject Node must defer; as it cannot transmit with enough power to reach its own Recipient (now 4 dB "further away") without also narrowing the Incumbent's margin below 15 dB.

The Second of two parts of this calculation is to systematize the computation of the four path loss values for each relevant combination of the I, R_I, and R_S nodes; so that each combination can be tested for Transmit or Deferral. Conceptually, the simplest process is to:

1. Select one of the 120 Nodes within Circle **B** as an I.
2. Select one of the 60 Nodes within an "**A**" radius surrounding this I as an R_I.
3. Select one of the 60 Nodes within Circle **A** as an R_S.
4. Compute the four path loss numbers for this Node combination.
5. Compute the Transmit/Defer ("T/D") qualifier and tally it as one or the other.
6. Repeat steps 3, 4 & 5 for each of the other 59 R_S Nodes within Circle **A**.
7. Repeat steps 2, 3, 4, 5, & 6 for each of the other 59 R_I Nodes within the "**A**" radius.
8. Repeat steps 1, 2, 3, 4, 5, 6, & 7 for each of the other 119 I Nodes within Circle **B**.

Simple analysis reveals that 60 x 60 x 120 = 432,000 separate computations will be needed. By reasons of symmetry, it is clear that Step 8 need be carried out only for I Nodes covering at most one quadrant of Figure 1, reducing the number of computations to 108,000. Further reductions may be possible, which will maintain the same ratio of Ts to Ds. An Excel or MathCad program could be written within a few hours to carry out the computations in a few minutes. It is extremely important to complete this calculation in order to determine whether or not implementation of this extensive WLAN system design change, in and of itself, will yield a significant improvement to overall throughput. The Author will either have this calculation complete or be ready to discuss it when he presents this Submission in Orlando 11 July.

E. Critical Issues

Although this section will not carry the largest impact, it is probably the most important for ultimate development of the best possible Standard. It is probably too difficult, in a short period of time, for one person to come up with all the best, exact answers. The best the Author could hope for is to come up with all the relevant questions.

The Issue of Public Policy

Before finalizing strategies to implement TDA, a legal opinion might be necessary. Notwithstanding the assurances (based on some legal foundations) given by McDonald in 94/79 hereinbefore discussed, that:

- The Band is available for a wide variety of uses;
- The FCC mandated spread spectrum to remove the need for etiquette;
- There won't be much interference between networks anyway.

We cannot automatically assume that government regulations, after being tested in the courts, will allow any one user to operate with total disregard to the welfare of other users which whom there is no privity (Foreign users). There may be a minimum required etiquette, and details of this requirement may crystallize during the next year or two as the Band becomes crowded and tempers begin to flare.

Of course, each manufacturer is concerned primarily with its own equipment and customers, and would at least like to approach TDA with a parochial hierarchy or "pecking order", with his own networks at the top. The "etiquette" or "protection priority" list might look like this:

1. This immediate (Native) network cell;
2. The remainder of this network;
3. Other Compatible networks owned by the same end user;
4. Compatible networks owned by different end users;
5. Incompatible but Compliant Networks;
6. Non-Compliant networks and other systems not operating under 802.11 standards, such as low speed signaling systems, high speed common carrier point to point T1 systems, and cordless telephones (Yes, cordless telephones will move up to the Band);
7. Unintentional industrial, scientific and medical radiators, such as microwave ovens.

As those who are likely to be listed as the writers of the Standard that will lead to the manufacture and sale of equipment which could be construed in a Court of Law as actually being designed to cause interference to other communication services on the Band, interference which may cause damage to the other users of this Band, we should be extremely careful about how the Standard is worded.

For instance, we should avoid if possible direct references in the Standard to transmitting during periods when other intentional radiators are operating, and merely provide for the infrastructure allowing each manufacturer to make his own tacit decisions. If such references, as the Author and many others strongly suspect, cannot be avoided, there needs to be some Deferral mechanism indicated for all Intentional Radiators. Inasmuch as Non Compliant equipment transmissions cannot in practice be identified by any other methods, it would appear that one component of the TDA criteria must be RF Power Detect.

Technological Issues

In Part C of this Submission, the Author suggests that the TDA process should involve a nontrivial combination of several variables. The experienced Radio LAN ("RLAN") Engineer Reader knows that the neither the detailed Standard specification nor RLAN equipment design taking best advantage of these variables will be at all trivial. Also In Part D, this method was explained. The Reader can now well imagine that the acquisition and management of such data involves much further investment in design time and resources. The present section attempts to state the issues which govern how these ideas might be incorporated into a Standard plan.

The Multivariate TDA MAC/PHY Interface

The present consensus of the PHY subgroup is that the TDA interface should behave like the RS-232 "CTS"; which consists of a single one-bit field, Clear or Busy. If Clear, ready packet fragments will be transmitted forthwith (but consistent with non-persistent and other backoff algorithms); if Busy, no transmission will be initiated. At present, the MAC Subgroup is expecting this single bit TDA. The three CCA papers (ob. cit.) already submitted to the PHY Subgroup propose criteria, processed solely by the PHY, for when this bit should say Busy and when it should say Clear. It is the Author's conviction that owing to vagaries of the indoor RF medium and the multiplicity of unrelated users, (in contradistinction to wired media or at least dedicated radio spectrum) the TDA process is so complicated and relies on so much field experience we do not have yet, in order to achieve the best possible network throughput it is advisable to consider presenting the appropriately programmed MAC Layer with an array of "raw" information and allow it to decide when to transmit. Furthermore, this decision needs to be made in conjunction with higher level information available only to the MAC.

More specifically, the MAC could be presented with a three octet state table containing the following information:

1. Five bits for a logarithmic RSSI scale
2. One bit for recovered clock (indicates the presence or absence of a Compatible LAN Node) in the midst of a transmission
3. Twelve bits for the transmission length field
4. Three bits for the transmitted power field
5. One bit for the Idle Pattern
6. Two spare bits

Each 10 microseconds, the existing values are clocked through a buffer. The MAC makes a TDA whenever necessary, based upon all these conditions plus information from upper layers, as a function of time.

List of Relevant Issues

1. The TDA decision depends jointly on three factors: 1) Needs of upper OSI layers; 2) Existence of a "Ready" Packet/Fragment; 3) Channel conditions (meaning all the conditions shown at the end of Part C, plus the "True" TDA plan).
2. Are the Channel Conditions separable from the Needs of the upper layers, or must a determination be made by a combination of both?
3. If the Channel Conditions are separable from the MAC functions, is the PHY capable of handling them in their present elaborate and convoluted form?
4. Including sending Network Maintenance packets?
5. If the Channel Conditions must be handled in the MAC, will they be PHY independent?
6. If the Channel Conditions are not PHY independent, does it matter, as long as the Node as a whole behaves per the Standard?
7. Or is the Standard intended to allow mixtures of different PHYs and MACs in the same OEM products?
8. Can the Node afford the battery drain, if the MAC participates in an ongoing fashion, gathering and processing the data needed for an elaborate TDA?
9. Not just when to transmit, but for how long (Fragment length)
10. If the PHY can make the entire Channel Conditions decision, separate from MAC functions, can it deliver one more interface field to the MAC containing its assessment of the maximum advisable Fragment length at this time?

F. Conclusions

If the Transmission Deferral Assessment methods explained in this Submission are incorporated into the Standard, the initial investment in engineering and time to market will be relatively large, but the marked superiority in RLAN throughput performance will pay off handsomely in the end..

Notes for Discussion