IEEE P802.11
802 LAN Access Method for Wireless Physical Medium

DATE: August 6, 1991 **REVISED:** August 26, 1991

TITLE: **ACCESS METHOD FOR CHANNELIZED SYSTEM USING DISTRIBUTED LOGIC AND NOT REQUIRING INFRASTRUCTURE**

AUTHOR: Chandos A. Rypinski,
Chief Technical Officer
LACE, Inc.
921 Transport Way
Petaluma, California 94954 USA

Telephone: 707 765 9627
Facsimile: 707 762 5328

SUMMARY

This medium access method uses one setup and nine (or N) data transfer channels in which any Station can transmit at any time on the setup channel without reference to whether or not signal is present. The channelization is assumed to be code-division within a spread spectrum modulation, but is not limited to this possibility.

The main assumption is that there is always a probability of a lost message or transmission from uncontrollable factors like path obstruction and multipath. Contention may be allowed within the system as long as its relative probability is less than or the same order of magnitude as other message loss mechanisms.

This access method is optimized for peer-to-peer communication without use of infrastructure. Infrastructure may be used as a means of providing communication for Station in the same network but not within radio range of each other, and to provide a means for each Station to reach destinations outside of the local network.

This access method is not suitable for a virtual circuit service.

This is the third of three access methods that have been developed all using the same message set and which can be characterized as follows:

1) With infrastructure, sequential use of one channel at all Access-points within one reuse group. (IEEE 802.11/91-19 and IEEE 802.11/91-95)
2) With infrastructure, sequential use of a common setup channel and parallel a number of data transfer channels derived by code-division or otherwise. (IEEE 802.11/91-97)
3) **Independent of infrastructure and without virtual circuit support, random contention use of a common setup channel and distributed channel selection for following parallel use of one of several data transfer channels.**

Table of Contents Page

TABLES

FIGURES

ACCESS METHOD FOR CHANNELIZED SYSTEM USING DISTRIBUTED LOGIC AND NOT REQUIRING INFRASTRUCTURE

OVERVIEW

The central premise is:

In a radio system, there is always a probability of a lost message or transmission from uncontrollable factors like path obstruction and multipath. Contention may be allowed within the system as long as its relative probability is less than or the same order of magnitude as other message loss mechanisms.

The major objectives of the chosen functional characteristics are to provide:

- a primary access method which is optimized for peer-to-peer communication using fully distributed logic and without use of infrastructure, and
- a backup optional infrastructure means of providing communication between peers in the same network but not within radio range of each other, and
- for the infrastructure to provide a means for each Station to reach destinations outside of the local network.
- given that spread spectrum modulation is used for better performance in multipath environments, to also realize the increase in the capacity of a given frequency space that is available from use of code division multiplexing.
- to use the same protocol message structure whether or not the infrastructure is secondary as in this proposal or primary where positive management of capacity is required.

The system design considers the following particular detail requirements:

a) when contention is possible, the shortest possible transmission lengths are used to minimize contention probability.

b) a systematic and rapid method of detecting and recovering from lost messages is used.

c) channel on-time is minimized for carrying a given traffic load.

d) no worst case delay requirements are imposed that are inconsistent with multiple transmission attempts.

e) an effective algorithm is used for selection of data transfer channel by distributed Station logic.

Highlights of the system are shown in Table I below:

Table I -- HIGHLIGHTS OF CONTENTION ACCESS CHANNELIZED SYSTEM

- Channelized into one SETUP and nine DATA TRANSFER channels.

- In the autonomous mode, the initiating Station may *REQUEST* access to an addressed Station at any time on the SETUP channel without regard to existing activity.

- The addressed Station responds with a *GRANT* on the SETUP channel selecting the DATA TRANSFER channel to be used.

- The initiating Station transmits the *PACKET DATA FRAME* on the selected DATA TRANSFER channel, and hears *ACK* from the addressed Station.

- Transmissions not acknowledged or otherwise unrecognized are repeated up to two additional transmissions.

- The addressed Station selects the DATA TRANSFER channel by incrementing the number of last channel assignment heard.

- With infrastructure, the protocol is the same except that when the address is recognized as a non-local Station or when the addressee does not send *GRANT*, an infrastructure Access-point responds and relays.

- Reduced power transmissions are used on the DATA TRANSFER channel only.

- To minimize activity time on the SETUP channel, short addressing is used for local Station-to-Station messages.

- Short addresses are assigned at registration when infrastructure is present, otherwise they are the two least significant octets of the long address.

- With 10% active time on the setup channel, the capacity is estimated at 550 packets per second at 1 Mb/s. For a 300 octet average packet length, a 16% loading of the data transfer channels results. It is unlikely that the data transfer channels would limit system capacity until there is more than 20% activity on the setup channel.

- Message set per IEEE 802.11/91-80, except Access-point format *GRANT* and *ACK* added to the Station message list.

Channelization

These objectives are attained by using one setup and nine data transfer channels in a nine channel reuse pattern for BSAs. The channels can be derived by code division with spread spectrum or by frequency division. The channelization provides the separation of a setup function with contention from data transfer with a very small possibility of contention.

Contention Detection

> **For this plan, the primary means of detecting transmissions impaired by contention or any other cause is lack of positive acknowledgment. No attempt is made to avoid contention by listening for an existing signal on the channel, because the presence of signal is an inconclusive indication that the next transmission will be harmful to existing channel use or unsuccessful on the new use.**

ARQ -- Automatically Requested Repeat

With an ARQ system providing up to three tries for any failed message, it is no longer necessary that every transmission be successful--a goal that is much harder to approximate in a radio system than in a closed cable system.

Provided that the cause of error is purely random, a message which has a 1-in-10 chance of being missed, will have a 1-in-100 chance of successful transmission with two tries and 1-in-1000 with three.

The first problem to be overcome is the correlation between consecutive transmissions. If there is an obstacle in radio path, it is unlikely that the second or any later transmission transmitted within 10 milliseconds will have any better result. This is a fundamental factor in the radio system design.

MECHANICS OF CONTENTION ON A CHANNELIZED SYSTEM

If a channel is active 1% of the time, the probability that the channel will be busy at a random instant when a new transmission might start is also 1%.

As long as the overlap probability is less than a few percent, there is no need to consider the consequences of the lost messages. The occurrence of contention increases the amount of channel time required to move a given amount of traffic by a like small percentage. At higher percentages it is necessary to consider the retry consequences.

At 10% contention probability the amount of air time used would be doubled assuming that each attempted use is tried again and successful. This is not negligible.

In a LAN, a Station desiring to send a packet will try again and again rather than give up and leave the channel. In a well loaded system the behavior of unsuccessful efforts to obtain access or send a message must be defined before analysis is possible.

The more manageable parameters in minimizing contention are the length of the contending transmissions and the frequency with which they occur.

Contention on the Setup Channel

When the second transmission of a contending pair is started, conventional theory (e.g. Kleinrock) would say that <u>both of the overlapping messages become useless. This is not true in a radio system</u>, where with two simultaneous transmissions, there is at least a 50% probability that the desired message is stronger at the addressed receiver. There may be at least a 15% chance that it is sufficiently stronger to be received successfully.

If the odds can be altered by saying that the probability of the interfering signal originating in the same area as the desired signal is small, a new basis must be found. This is inherent in a signaling channel shared by 9 or 16 coverage areas. There is then less than 11% probability (now rounded off to 10%) that the interfering signal originates within the coverage area of the desired signal where there is a good chance of high enough level to interfere. In this case, the probability of detectable signals being interfering signals might drop from 85% to 8.5% of the cases.

> Because most of the interfering signals probably come from greater distances, there might be a better than 90% probability of a transmission being successful even though another Station in the same 9-BSA group is transmitting at the same time.

Contention on the Data Transfer Channel

For the 9-BSA model assumed where each BSA has an assigned data transfer channel, the setup function is allotted 10% of the spectrum space used. The data transfer traffic is divided over 9 channels which can coincide with the BSA (but there are other possibilities for transfer channel assignment). <u>The selection of the data transfer channel without a central or common access manager cannot assure that there is no damaging interference on the selected channel</u>. Algorithms, more ingenious than location associated indices will result in substantial improvements in realizable capacity.

PROTOCOL FOR SETUP AND DATA TRANSFER

The system logic may be based on usually rather than assuredly successful functions. Any Station may use the setup channel at any time to initiate a transfer, and usually this attempt will be non-interfering with any other; but if it is, there is a retry mechanism.

The addressed Station will respond to a request, usually sending a *GRANT* (115 = 015 except Station-originated) message that includes nomination of the channel on which the initiating Station should send the message.

Usually, the nominated data transfer channel will be clear, but not certainly.

Autonomous (No Infrastructure) Case

The distributed algorithm for data channel selection depends on the fact that each Station listens to the setup channel continuously. The current next-data-transfer-channel-to-be-used is the last one assigned incremented by one.

The requesting Station hearing a *GRANT* message from the addressee moves to the nominated data transfer channel and sends the packet. If successful, the addressed Station sends *ACK* (110).

If acknowledgment is not received, there are two possible points of failure that are not distinguishable to the originator suggesting the desirability of an acknowledgement of an acknowledgement.

It is also useful for the requesting Station to announce (broadcast) the end of use of a data transfer channel on the setup channel; and this would serve the second function enabling the addressed Station to ask for a resend on failed acknowledgement. The second *ACK* (111) would use the Access-point format (011) containing the CHL field.

This procedure offers good probability of successful transfer, but it does not assure it.

With Minimal Infrastructure Assistance Case

With infrastructure, the simplest case follows identical steps except for the means of deciding which data transfer channel to select.

Each access point may broadcast its existence with a message at one second intervals. The Station then uses the data transfer channel associated with the Access-point which it currently prefers. Further refinements of this logic may avoid the need for a Station to measure signal level to choose between multiple access points being received.

Intelligent and Coverage-extending Infrastructure

When a Station requests and the addressed Station does not respond, there is nothing the originator can do but try again. This is useless if the addressed Station is out-of-range or OFF. Too much repeat effort would spend system capacity and battery for without useful result.

It is possible for the infrastructure to know whether Stations are ON and which Access-points are providing satisfactory coverage of each

Station, and this information is vital for an ESA type system.

When a Station requests, it may be possible for the infrastructure to know whether communication directly to the addressed Station is possible. If it is a known direct-impossible address, the infrastructure Access-point can issue the *GRANT* (015) and accept the message for repetition slightly delayed but simultaneously on another Access-point or upon completion on the same Access-point.

If the possibility of direct communication is not known, the infrastructure can wait to see if the addressed Station sends *GRANT* (115) and if not, the infrastructure can then *GRANT* (015) and receive as described above.

The infrastructure can respond directly to messages addressed outside of the local area network depending upon servers and/or bridges for outside relay.

Something that the infrastructure can do with near certainty is know which data transfer channels are currently in use and therefore unavailable. With infrastructure and relay, it is not mandatory that there be a geographic association with individual channel numbers.

If the Stations operate autonomously there is no assurance that the data transfer channel selected is available. If the channel to be used is decided by reception of an Access-point broadcast, it is possible for any Station to know that that channel is busy for defined time period by listening to the setup channel provided that one of the two Stations involved is sufficiently close to be heard directly. There is no certainty, only probability, that the appropriate messages will be heard by any particular Station.

ASYNCHRONOUS MESSAGE-BASED CONTROL PROTOCOL

The protocol is called asynchronous because no use of time-framing or regular periodic slots is made. Except for the initiating *INVITATION-TO-REQUEST/REGISTER/POLL* messages, each following message is transmitted when the prerequisite message has been completed.

The message set used has been defined for infrastructure associated systems as given in IEEE 802.11/91-80. A subset is used in this plan.

The main differences are additions to accommodate operation without infrastructure. A provision is made for Stations to use the format of normally Access-point originated messages 011 and 015 retyped 111 and 115.

There is an implied time interval structure that depends upon the length and format of the messages of which the transfer functions are composed. The transmission of a packet to an in-range peer requires the steps shown in Table II on the following page.

The saturated transfer rate could be reached if all messages lengths were equal to the time between consecutive setups for one data transfer channel which is about 180 octets with short addressing.

The time required is the same for Station or Access-point originated traffic with or without infrastructure. For the case where the infrastructure waits to hear if the addressed Station responds, another 100 μseconds is added to the time required.

Comments on Short/Long Addressing

The time cost of always using long addresses amounts to 1/3rd of the setup channel capacity. This is inducement for retaining short addressing within local groups. The default short address is the two least significant octets of the LAN address. For those Stations that normally work together, and are close enough to communicate without infrastructure assistance, alternative short addresses could be entered manually in case of duplication.

With infrastructure, the assignment of non-duplicating short address is easily accomplished as part of a necessary registration function.

Long addresses are clearly required for addressees outside of the local network or for foreign users within it.

For traffic estimates, it is assumed that all traffic uses short addresses which is reasonable when there is no communication outside for which infrastructure is used.

Table II – TIME INTERVALS FOR A CHANNEL DATA RATE OF 1 Mb/s
DISTRIBUTED LOGIC ACCESS METHOD

		octets		µseconds	
1a)	*REQUEST*-long address (108)	23		194	
2)	*GRANT* (115)	9		82	
	Time on setup channel:			276	
1b)	*REQUEST*-short address (108)	11		98	
2)	*GRANT* (115)	9		82	
	Time on setup channel:			180	
3)	*PACKET DATA FRAME* (114)	10	+PDU	90	+ PDU
4)	*ACK* (011)	8		74	
	Time on data transfer channel:			164	+ PDU
5	*ACK*-channel release (111)	8		74	
Time required for 9 setups--short address:				1620	
Time required for 9 setups--long address:				2484	
Setup Capacity Limited:				Number	
Transfers per second saturated--long address:				≈ 3600	
Transfers per second saturated--short address:				≈ 5500	

Capacity Estimates

For average packet lengths shorter than some value above 180 octets, the system capacity is limited by the setup channel loading. The contention access method must be considered for one channel common to nine BSAs, and it becomes doubtful if more than 10% of uses result in contention.

The amount of activity required to reach this limit is not obvious because a fraction of the simultaneous usage will not result in transmission failures. Also some of the simultaneous usage will result in two lost transmissions rather than only one. For the moment, it is assumed that these two factors are offsetting.

> For the above assumptions: 10% activity on the setup channel is 550 packets per second at 1 Mb/s. For a 300 octet average packet length, a 16% loading of the data transfer channels would result. It is unlikely that the data transfer channels would limit system capacity until there is more than 20% activity on the setup channel.

POWER LEVEL AND CONTROL

In any large system it is desirable to limit transmitter power to that necessary. It is inevitable that *REQUEST* be transmitted at the maximum power available. It is desirable for *GRANT* also to enable a power reducing algorithm to start from a known level, and because that message has a broadcast significance to other Stations in the network.

After setup, both addresser and addressee know how much signal margin is available and can reduce power accordingly for the transmissions on the data transfer channel.

This function creates a need for fast signal level measurement in the Station receiver which is not present in the plans primarily depending on infrastructure. For this reason, implementation of power control is considered unsuited to non-infrastructure system plans.

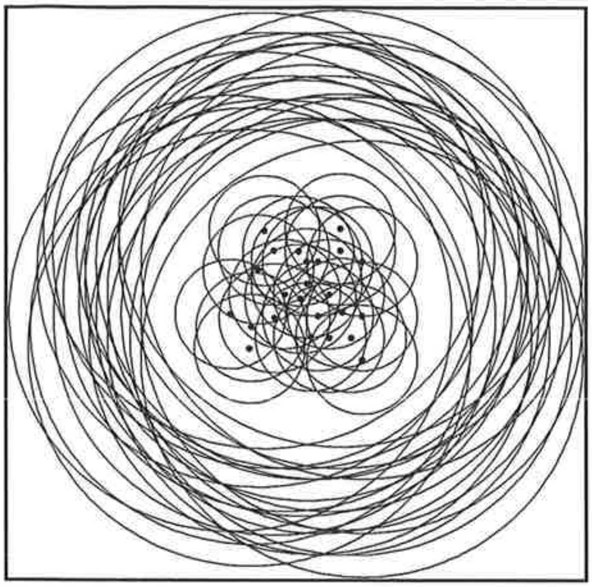

Figure 1 Semi-random location pattern of 24 Stations showing service and detectability range circles.

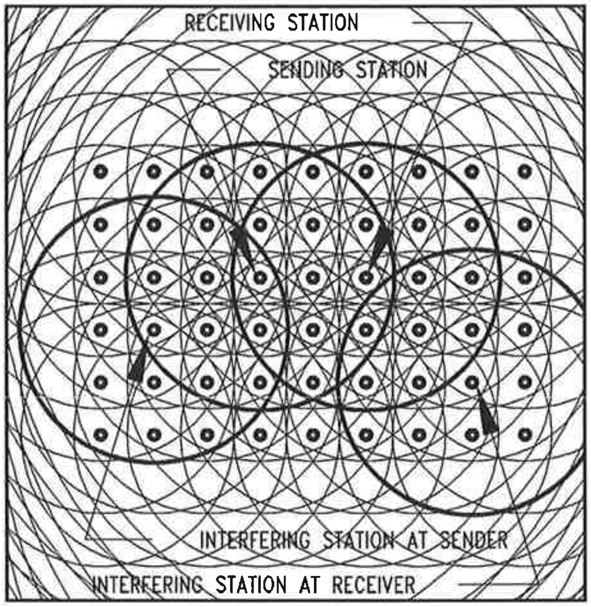
RECEIVING STATION
SENDING STATION
INTERFERING STATION AT SENDER
INTERFERING STATION AT RECEIVER

Figure 2 Pattern of 54 regularly spaced Stations showing service and detectability range circles.

INTERFERENCE PROBABILITY BETWEEN STATIONS AND ITS EVALUATION

The probability of interference between Stations is not often analyzed because it is difficult to avoid unrealistic assumptions about their arrangement. However, some important notions about interference may be found. As shown in Figure 1, a number of randomly located Stations may be thought of as having:

a) A service range with 95% probability of a message being correctly received, and
b) An interference range with a 5% probability of the signal being completely decodable at a receiving Station which is possibly 4X the service range.

The service range is the smaller circle and the interference range the larger. From this Figure a few conclusions can be reached:

c) All of the Stations have some probability of being receivable at any other, and
d) Only a few of the other Stations are within the service range of any particular Station, and
e) Only a fraction of the receivable Stations will be close enough to prevent communication between Stations within the service range, and
f) the common service area for a group of Stations is much smaller than the service area of any one.

Figure 2 is similar to Figure 1, but the arrangement of Stations is regular. This makes it easier to interpret principles that may exist even though most real situations will be irregular.

The point of this Figure is to show that while a potentially interfering signal is probably detectable at both transmitting and receiving Stations, the following points may apply:

g) a signal received at a high enough level to be interfering at the transmitting Station is not necessarily interfering at the receiving Station, and
h) the receiving Station is in a better position to evaluate interference than the transmitting Station, and
i) the coverage and interference ranges of a Station are not fixed values since it is signal-to-interference ratio that determines receivability of a desired signal.

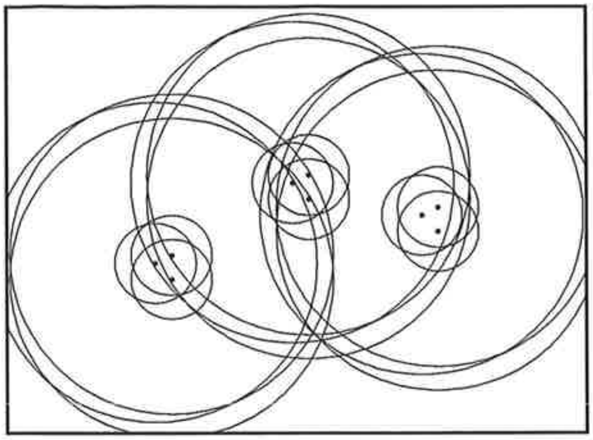

Figure 3 Separated clusters of Stations with service and detectability range circles.

In Figure 3, separated clusters of three Stations are shown similarly. It is easier to see the diminished common coverage area relative to the area of any one Station. It is also notable that the precise location of the Stations relative to each other is much more important for coverage than it is for interference.

CONCLUSIONS AND EVALUATION OF PEER-TO-PEER NETWORKS

To fairly evaluate peer-to-peer networks, it is essential to have a detailed access method that spells out the initiation, transport and termination of a packet transfer. The first part of this paper is an effort to provide a rationale and to attempt the best possible implementation of a contention access method.

Anticipated Conclusions

From the beginning, some of the conclusions have been anticipated as follows:

1) The observation of channel activity, either by carrier-present sensing or by detection of valid data transmission, is an undesirable method of enabling transmission.
2) If a Station is going to transmit at random but infrequent instants, there is more to be lost than gained by waiting for a clear channel.
3) The usability of contention access is based on the following points:
   a) separation of the contention possibility on setup from data transfer requiring a channelized system
   b) less than 25% air-time utilization
   c) a good ARQ (auto repeat) algorithm for recovering from missed messages

Absence of Access-points

For a work group of several Stations, the area of interoperability is the common area of their respective individual coverages. The interference potential of this group is greater than the interference range of any one Station which effects the possibility of similar groups at a distance. The only mitigating factor is that if the usage of all Stations within interfering distance of each other is a very small fraction of the available air-time, operation may be satisfactory.

The presence of an aggregate background interference from nearby systems will result in a shrinkage in coverage from a range limited by detectability to a range limited by the background interference.

There are so many probabilities in tandem in estimating the serviceability of a direct peer-to-peer network that there are few examples of realistic modeling and little means for predicting capacity and performance.

Some of the gains from use of Access-points are as follows:

4) All Stations within range of a common Access-point are in range of each other
5) The limits of a Basic Service Area are definable corresponding to that of an Access-point
6) An Access-point with ceiling-height antenna has greater range than a Station with table-height antenna
7) Interference can be reduced by location of Access-point relative to natural obstacles and interferers

For a non-infrastructure peer-to-peer only LAN service, it is possible to use a channelized system with a contention setup and reduced contention data transfer channels. Such a system might be used as a subset of an on-demand infrastructure system. For a combined system, it is desirable and possible to use a common message set.

The time, spectrum and power utilization of an infrastructure system will be much better than for a peer-to-peer only system.