IEEE P802.11

Wireless Access Method and Physical Layer Specification

DATE: October 30, 1991 **REVISED:** 0

TITLE: **ARCHITECTURAL CONSIDERATIONS FOR LARGE SCALE WIRELESS PERSONAL VOICE-DATA NETWORKS**

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

Telephone: 707 765 9627
Facsimile: 707 762 5328

SUMMARY

The traditional criteria for spectrum efficiency, cost and degrees of geographical distribution of functions must be considerably revised for the context of high capacity premises area networks, and many new criteria must be introduced for which there are either narrow or insufficient precedents. An effort is made to identify and define these criteria converting them into constraints and preferred forms for the system architecture.

CONCLUSIONS

Considering the architectural alternatives and available technology tools, and within the context of premises area network size, the following conclusions are reached on the choices necessary for a satisfactory total outcome:

1) An entirely burst mode transmission format (like ATM cells) must be used with all necessary processing information for the receiving point contained in a header of each burst.
2) Local shared mediums should be used serially and asynchronously at full rate.
3) The distance between Station and access controller measured in round-trip processing and propagation time should be short compared with the shortest time-duration burst commonly carried on the medium.
4) Various densities of capacity can be produced by scaling of physical medium signaling rate and the coverage area of each Access-point; however, as the coverage per Access-point increases, the total capacity per Access-point will decrease exponentially.
5) Each bridge to an 802 LAN or other backbone must be common to a number of Access-points sufficiently large to disengage the significance of the Access-point used by a Station or Station-associated bridge to external spanning tree bridges or to source-routed Station algorithms.

Table of Contents Page

Tables and Figures

ARCHITECTURAL CONSIDERATIONS FOR LARGE SCALE WIRELESS PERSONAL VOICE-DATA NETWORKS

By Chandos A. Rypinski

SUMMARY

The traditional criteria for spectrum efficiency, cost and degrees of geographical distribution of functions must be considerably revised for the context of high capacity premises area networks, and many new criteria must be introduced for which there are either narrow or insufficient precedents. An effort is made to identify and define these criteria converting them into constraints and preferred forms for the system architecture.

Given

The opening assumptions for the wireless network are:

1. It must be structurally capable of providing all of the telecom and computer data services now available at desktops.
2. The size, scale and capacity of the total service rendered must be a significant proportion of that provided by cable.
3. The equipment size, cost, reliability and maintainability must be comparable to present alternatives.
4. A possible solution exists.

The given universality requires interworking with IEEE 802 Standard LANs providing connectionless services, wide area store-and-forward internets, wide area SMDS integrated facilities and other existing isochronous connection-type services.

This problem is considered in the context of defining a Standard air-interface within IEEE 802.11, which requires consideration of all the functions on the Station and Infrastructure sides of that interface.

Conclusions

Considering the architectural alternatives and available technology tools, and within the context of premises area network size, the following conclusions are reached on the choices necessary for a satisfactory total outcome:

1) An entirely burst mode transmission format (like ATM cells) must be used with all necessary processing information for the receiving point contained in a header of each burst.

2) Local shared mediums should be used serially and asynchronously at full rate.

3) The distance between Station and access controller measured in round-trip processing and propagation time should be short compared with the shortest time-duration burst commonly carried on the medium.

4) Various densities of capacity can be produced by scaling of physical medium signaling rate and the coverage area of each Access-point; however, as the coverage per Access-point increases, the total capacity per Access-point will decrease exponentially.

5) Each bridge to an 802 LAN or other backbone must be common to a number of Access-points sufficiently large to disengage the significance of the Access-point used by a Station or Station-associated bridge to external spanning tree bridges or to source-routed Station algorithms.

The implications of 1) to 3) are that the use of channelized media must be avoided whether division by time, code or frequency is used. As soon as the needs for fast setup, bandwidth-on-demand and use of redundant transmission paths is considered, there is no acceptable alternative.

A fine line is drawn between time-division multiplexing which is essential and time-division channelization which is commonplace in telecom. The channelized medium is defined to include all framed systems with a period of 125 μsec.

The background and reasoning by which these rules were arrived at is the substance of this contribution.

A simplified graphic representation of such a system is shown in Figures 1 and 4.

OPTIMIZATION CRITERIA FOR CHOICES

The better possibilities are groups of choices. The optimization of any one or two criteria does not have value. The cost of Station, infrastructure and maintenance together are a primary criteria; but separately, they are only important factors.

Physical Medium Criteria

The most arguable factors are medium signaling rate, modulation technique and spectrum efficiency. If the medium signaling rate is the initial choice, then all other factors are variations in either the achievable reach limit from time dispersion or the occupied bandwidth.

> Roughly, the reach-rate product is a constant which decreases as the complexity and number of levels in the modulation increases.

The gain of complex modulations from increased symbol length is offset by decreased noise and crosstalk tolerance. More power is required to transfer at a given information rate with complex modulations.

Radio Bandwidth, Modulation and Rate Tradeoffs

If the amount of spectrum is taken as a given, the available bandwidth puts a ceiling on the signaling rate as a function of modulation method. If the bandwidth is divided into channels, the peak data rate available is also divided provided that it is not already limited by the time-dispersion for the assumed reach requirement.

It is necessary to evaluate radio modulation as means of expanding the transmission limits.

> A radio modulation that is less spectrum efficient in a narrow sense may be the most robust against like-signal interference and time dispersion in the medium leading to the best spectrum efficiency in a system plan.

> It follows that the modulations utilizing spectrum at better than 1 bit/Hz are not only more costly, but they are sufficiently impaired in interference tolerance that there is a loss in spectrum efficiency.

Regulatory Spectrum Use Limitations

Choices are greatly influenced by concepts of what the physical medium limitations are. Radio systems are most sensitive to the assumed regulatory and frequency allocation environment in which they will be operated. Inexperienced individuals will take the present radio regulations as absolute; however,there is elasticity in some areas and possible major change in others which are known to more experienced spectrum users.

A common difficulty is the use of frequency hopping for no other reason than rule compliance. More attention should be given to the intent of limiting power density on any given channel.

The Nature of the Radio Channel

As established by extensive experimental data, there is no way to make a radio path work with the approximation of certainty of contained mediums like optical fiber or even unshielded twisted pair. Good modulation and coding technology will make it less uncertain and no more.

> A radio system only has a statistical probability of successful transfer. This uncertainty applies not only to success but to failure and the appearance of foreign signals. An access protocol designed for near certain systems is unlikely to be suitable for a probabilistic medium.

Shown in Figure 2 is a qualitative description of radio range with omni-directional antennas. As distance increases, the probability of a successful data transfer decreases. The interference range is several times the service range. If stations are spaced according their reliable service range, there will be gross coverage overlap.

Interaction of Choices

There is direct interaction between access methods, compatibilities and radio modulation technique. A sufficient result cannot be assembled by taking existing pieces from here and there particularly for an access protocol. This is because of a narrow criteria set for optimization performed in the choice of nearly all existing physical mediums and access methods accepted as Standards.

Service Requirements

The performance requirements of the services normally provided must be negligibly degraded from the same functions on wired mediums.

The extent and capacity for these services must equal that on the preponderance of desktops when the system is maximally scaled upward. The standard chosen shall not technically preclude replacing all wiring in moderate usage intensity environments, though this may be a rare case. Also, when the system is scaled to replace only a small fraction of the wired services, the economic penalty for the greater capabilities must not burden small users and battery-powered Stations.

These requirements must not preclude operation by spontaneous, autonomous groups where no infrastructure exists.

CAPACITY AND RADIO SYSTEM TOPOLOGY

The considerations involved in choosing signaling rate, reach and channelization may be most easily explained by pointing out the alternatives for a system providing wireless access for a floor area of 100,000 meters². If all Stations were served by wireless, there might be 6,250 Stations or one per 16 m². Possibly only 20% of these, 1,250 Stations, must really be wireless.

The available radio frequency spectrum is sufficient to provide 2 Mbits/sec throughput at 2.45 GHz using spread-spectrum at 31 chips bit.

First-Try Dimensioning

The first step is to estimated the aggregate average traffic for 1250 Station using both duplex telephony and packet data. The users are those with some particular benefit from wireless. It is assumed that high intensity telephone users (e.g. purchasing departments) are excluded from the system.

The load estimate must be adjusted upward to allow for overhead (50% of air-time) and allowable maximum utilization (20% of available channel time). If each Access-point has a capacity of 2 Mb/s, then 48 would be required assuming that they could all operate independently on a common radio frequency channel.

The utilization dimension assumes that the average air-time for Stations and Access-points combined is 20% so they are not all ON at the same time, but most would be on at frequent intervals. This factor is not enough to deal with the overlapping coverage and resulting interference that is inevitable when all are on a common channel.

TABLE I – USAGE ESTIMATE
(1,250 Stations)

Type Use	Time-Usage min/stn/hour	Digital Usage Mbits/stn/hour	Total Usage Mbits/time	
Voice @ 32 Kb/s:	2	7.68		
Data:		20		
Voice+data:		27.7	34,600	Mb/hr
			9.61	Mb/sec
After overhead/utilization adjustment, **total aggregate system load:**			**96.1**	**Mb/sec**

TABLE II–SYSTEM SCALE FACTORS
(100,000 m^2)

# Access-points	Area m^2/ Access-point	Length side □ coverage	Distance to corner	# Stns/ Access-point
48	2,083	45.6	32.3	26
192	521	22.8	16.2	6.5

Options for Reducing Co-channel Interference

A first solution is to reduce the demand in relation to the number of Access-points. If the probability of overlap interference can be made sufficiently small, then the automatic retransmission feature of the access protocol might provide a sufficient recovery mechanism.

If the same traffic is spread over four times as many Access-points, the probability of interference is reduced because any one Access-point is transmitting or receiving only 5% instead of 20% of the time. This is a great improvement, but still not sufficient

A second solution is to channelize the system. It is possible to design an access method using one setup channel and nine data transfer channels. Each access-point uses an assigned data transfer channel so that there is a reuse factor of 9 with respect to co-channel interference. This channelization might be accomplished with the code-division feature of the spread spectrum modulation. There are technical limitations which means that the channel separation will work most of the time but not absolutely all the time.

If the channelization is done with frequency-division, then 10 times as much frequency space would be required. This possibility can only be considered with use of a narrow band modulation, and then the resistance to multi-path caused time dispersion would be much less.

A third solution is a structure use of channel time at Access-points combined with use of a larger number of Access-points.

Reduced and Structured Active Time

Suppose that the number of Access-points be increased by four times with the proportions shown on the second data line of Table II above.

A new factor is now introduced: Groups of four access-points are used sequentially, and there is then no interference between members of the group because they are not active at the same time. Each Access-point in the group would be active 5% of the time to carry the same traffic.

A further cooperative active time between each group and its neighbors can be added since the required active time for 15 Access-points is still only 80%.

Without intergroup coordination, interference can be reduced to something approaching 25% of what it would otherwise be (for this example, 1.25%). The rule that is inferred is:

> If many Access-points are each randomly active for small fractions of the time, the co-channel interference probability can be further reduced by use of structured patterns for the active time.

Increased Medium Signaling Rate

A much less obvious factor is that the shorter range implies a higher possible data rate. With half the required reach, the possible medium signaling rate might double if it is limited by time-dispersion or power. To accept the gain, it would be necessary to double the bandwidth or find some other equivalent improvement.

It is possible that the code-space used for channelization could alternatively be used to multiply the number of bits per symbol. Doubling the rate would then use the same bandwidth previously given for a code-division channelized spread spectrum system.

> The loss in capacity from sequentializing rather than parallel transmission and reception within a small group is recovered entirely or partially from the increased data rate possible with shorter reach.

In the example given, the load was first carried by 48 and then 192 Access-points holding data rate constant. The data rate and system capacity could be doubled relative to the time-dispersion range limit.

Benefits of Shortened Radio Range

Improvement in radio performance and easing of fading and time-dispersion problems occurs as the necessary range is reduced. In cluttered/mildly-obstructed environments, the probability of path obstruction is much less within 10 meters of the Access-point but is increasingly probable as the distance increases.

> The amount of power margin required to overcome obstruction loss decreases at short range reducing the ratio of the interference range to the service range.

Another form of the same statement is that the reach-rate product increases at shorter ranges because there is less loss from the interference of more distant co-channel transmitters. In the above example, the gain in interference probability is better than the obvious geometric ratios.

MOBILITY, FADING AND RADIO SYSTEM TOPOLOGY

In current mobile telephony and since early air-ground radio telephone systems, there is a "handoff" function to maintain call continuity as a Station passes from one coverage to the next. An "outage" of a second or less is tolerated on a voice circuit when this occurs. With a geographically large network it is costly or impossible to maintain redundant paths just to provide absolute continuity.

With packetized, short-reach radio systems, continuity does not exist in the medium, though it may required at the perimeter of the network. Multiple usable paths exist between Station and Access-point, and for any particular Station, a currently usable path can disappear in milliseconds while another is still satisfactory.

In the technology of contained mediums, path redundancy is a nuisance. One reason is because the network has no way to deal with multiple copies of the same message except as interference. In addition, the current routing algorithms in IEEE 802.1 have no speedy way of adapting to frequent movement of Stations between networks (or Access-points).

> A suitable design for the radio system and access method would turn path redundancy into an asset with the same value as N-port space diversity.

It is necessary to reconcile this observation with the existing routing algorithms for local area networks (e.g. "spanning tree" and "source routing").

Radio Access-points On A LAN Backbone

Many experienced LAN designers and users see a radio Access-point as a tap on an 802 or FDDI backbone. The resulting network might be configured as shown in Figure 3. Each Access-point has the radio physical medium and access method connected to the backbone LAN through a Bridge.

The Bridging Function

Upon receipt of a packet from a wireless Station, the local access method would deal with the case where the address is another Station on the same Access-point, and the Bridge would put externally addressed messages on the backbone LAN.

For messages heard on the backbone LAN, the Bridge would compare each received destination address with those known to be active on the associated Access-point forwarding the selected messages to the wireless access control.

This is straight-forward unless someone attaches another bridge to another network at one of the wireless Stations. At that point, recognition of messages that should be forwarded through the Station-bridge includes Stations that are not immediately known to the Access-point Bridge.

The routing algorithms address the need for bridges (or source routing Stations) to "learn" the forwarding patterns for a changing active Station data base. This is accomplished with broadcast inquiry and response messages containing nodal addresses that is triggered by the appearance of a new source address on a network.

Variable Location Stations

Suppose that a regular wireless Station goes from the coverage of one Access-point to another--possibly because someone walks between that Station antenna and the current Access-point or because it is a mobile Station. Until that Station initiates a new message, or is addressed and found unreachable from the previously current Access-point, the current Bridge does not know that the Station has moved. Even worse, this type of information may be stored in many bridges throughout an extended network.

Only with polling can change of location or status of Stations be found when the Station has no traffic. Without polling, there is a "loophole" state for messages toward moving Stations.

> With a bridge-per-Access-point topology, appreciable time and message activity must occur after a Station moves to the coverage of a different Access-point before all of the location and route records catch up. Further, this topology considers redundant radio paths a problem.

Hub Controller Network Topology

As shown in Figure 4, a number of radio Access-points can be connected to a common Hub Controller which is then bridged to any other Network. This topology becomes increasingly attractive as the reach and coverage of one Access-point diminishes.

> The primary value is that outside routers throughout the connected network need not be aware of any further addressing detail than that of the Hub Bridge regardless of which of the connected Access-points is used by a particular Station.

> The next value is that the function of using redundant transmission paths constructively is self-contained within the Hub Controller, and it does not affect routing in bridged networks.

There are additional values in the implementation details. In particular, there is a designer's choice in partitioning function location between the "ceiling mounted" Access-point and the Hub Controller.

"Bus-in-a-Box"

This term (previously used by A. Acampora) implies that a backplane with per-port electronics in the slots provides many of the same communication paths that would otherwise move through the backbone LAN in the previous topology.

> It is better to have the common bus that carries the traffic between Access-points run the length of one equipment box, than of the served premises area. For the short length bus, parallel and simultaneous events can be handled.

Considering the transmission line interconnecting each Access-point and the common Hub, the traffic carried in this case is only that of one Access-point. In the preceding case (Bridge-per-Access-point), the line carries all inter-Access-point traffic for all of the Access-points and all traffic in and out of the LAN. Since the backbone LAN has its own traffic limit, there can be a further capacity bottleneck.

With the bus in the box, the backplane may be made with as much capacity as it needs.

Processing of Multiple Received Message Copies

The nature of the system design doesn't permit much time skew between messages from one Station received on multiple Access-points. Each copy can be screened independently for CRC error indication. Any valid copy may be used. It does not matter which Access-point received the message once it is put on the backplane. With sequential use of Access-points, it will be the only message at that time within the sequential group.

Use of Redundant Paths Transmitting to Stations

The first try toward a Station would use the previously selected preferred Access-point. If this transmission is not acknowledged, then repeat transmissions are used which may come from different Access-points.

The protocol implications are that immediate acknowledgement is an essential feature, and that there may irregular time of receipt. Each segment of segmented packets should be numbered so that a Station can detect lost segments and assure the proper order of reassembly. Since segments may be transmitted at different times from different Access-points, this is an important point.

Simultaneous Virtual Circuits and LAN Packet

A Station could be supporting simultaneous virtual circuits and one packet-in-process simultaneously. It is necessary that the protocol provide markers and identifications sufficient that simultaneous activities within one Station do not confuse the processing at the receiving point.

ANALYSIS

1) An entirely burst mode transmission format (like ATM cells) must be used with all necessary processing information for the receiving point contained in a header of each burst.

The burst mode is inherently required because the time used to transmit information should be large compared with the time required to propagate, synchronize and identify.

The receiving point should learn how to process the burst payload from information in the header and not the time of arrival. If time slots are used to identify the associated connection or packet, this becomes a form of secondary local addressing on which a negotiation and agreement is necessary between sending and receiving points. The inherent delay in this process is not acceptable for LAN type communication even though overhead slot addressing would save channel time for long packets or virtual circuits.

Header-based processing is inherently necessary for the following reasons:

a) to avoid negotiation between transmitting and receiving points on the slot address to be used, and
b) to cope with variable access-delay, and
b) to use beneficially path redundancy in the medium, and
c) to screen unexpected foreign transmissions.

2) Local shared mediums should be used serially and asynchronously at full rate to deal with overlapping coverage.

The serial use of Access-points is a conclusive way to avoid interference effect within a "reuse-distance-defined" group. This is not the case for CDMA and for adjacent channels in FDM.

The aggregate capacity of a group of parallel channels is about the same as one serial channel occupying the same frequency space. The single channel is simpler protocol and faster response time. In addition, the radio is probably less costly because it does not need a fast switching channel synthesizer and has a larger fractional bandwidth reducing the required accuracy of filters and frequency control elements.

The wideband channel is advantageous for a spread-spectrum modulation with higher resistance to channel time-dispersion. The severe dynamic range requirement is avoided that results from code division channelization. It is also a simplification not to have channel switched codes.

3) The distance between Station and access controller measured in round-trip processing and propagation time should be short compared with the shortest time-duration burst commonly carried on the medium.

This criteria is not commonly stated. If the shortest message at 10 Mbits/sec is 12 octets, it only takes 9.6 μsec of airtime. The transmission time from the Station to the Controller and back might be as long as 4 μsec--a large percentage of the message time.

These short messages are inevitable for the invitations, grants, acknowledgements and for many payloads. Padding out the minimum message length does not avoid the time loss.

This consideration sets a limit on the usefulness of very high data rates from an efficiency viewpoint, and on the distance between Hub Controller and Access-points. 100 meters of UTTP is .7 μsec long (v.p.=50%) and it is probably a good assumption for limiting length.

4) Various densities of capacity can be produced by scaling of physical medium signaling rate and the coverage area of each Access-point; however, as the coverage per Access-point increases, the total capacity per Access-point will decrease exponentially.

If constant data rate is assumed, halving the range of an Access-point will quadruple the number required and the aggregate capacity of the system (square rule). If the permissible data rate is increased as the range is shortend, quadrupling the number of Access-points would increase the total capacity 8-times (cube rule).

In the examples, proportions have been suggested that are thought to be representative of real systems.

5) Each bridge to an 802 LAN or other backbone must be common to a number of Access-points sufficiently large to disengage the significance of the Access-point used by a Station or Station-associated bridge to external spanning tree bridges or to source routed Station algorithms.

The alternative is bridge-per-access-point. A station moving from one Access-point to another triggers rerouting and potentially a large number of service messages. If this occurs less frequently than once per hour, perhaps it is tolerable. For moving stations and smaller coverage Access-points, it could create message storms. This possibility affects outside equipment to an unpredictable degree.

6) Providing connection-type and connectionless services from a common infrastructure is compatible with the above conclusions.

The demonstration of this proposition requires examination of details of specific access protocols with respect to capability for capacity partition management and worst case access delay. That is beyond the scope of this paper but is believed to be true though not now shown.

If these generic conclusions were accepted as valid constraints by the 802.11 Committee, then the effort to develop a Standard could be considerably more focused than at present.

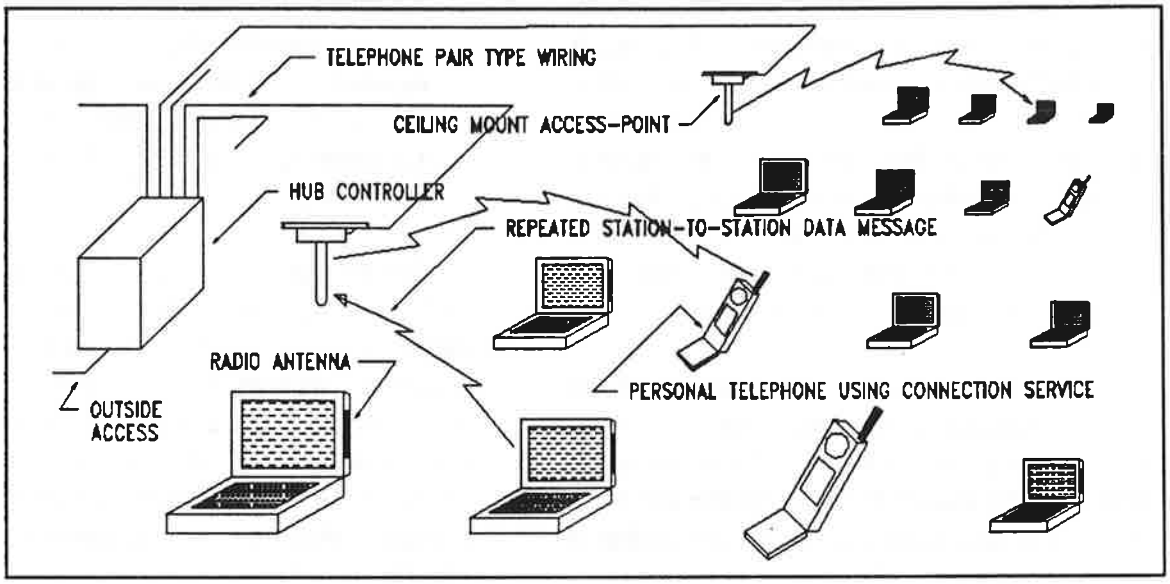

Figure 1 UBIQUITOUS WIRELESS ACCESS SHOWING INFRASTRUCTURE

This premises area radio access environment enables the portable computers to reach and be reached by outside entities using not only well known low-speed modem interfaces (e.g. X.25 protocol at 2.4 Kbits/sec) but also newer, faster public network transports at 64, 384 and 1.536 Mbits/sec. Packet LAN capabilities are extended to any other Station in the network, and is not limited to those within immediate radio range.

Pocket telephones are an additional use for the connection-type services provided.

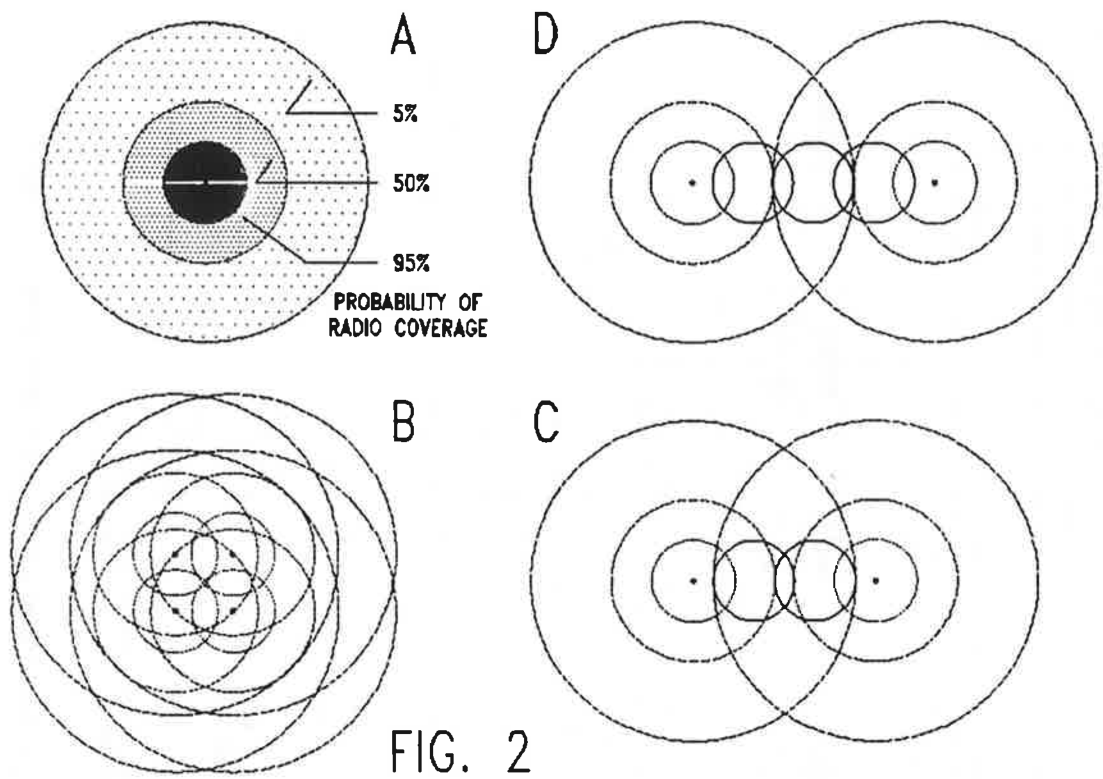

FIG. 2

Figure 2 RADIO SERVICE AND INTERFERENCE RANGE

Figure A is a qualitative illustration of radio coverage in a environment of moderate path obstruction. Coverage occurs in optical shadows as a result of reflections from other surrounding objects. On the average, the probability of adequate signal level and quality decreases with distance, however in any real system, the level and quality depends much more on the kind of obstructions and reflectors that are in or near the straight-line radio path.

The occasional unobstructed path may allow the signal to be used or create interference at a great distance. The normal design margins for solid coverage provide adequate signal at four-times design range when there is no obstruction.

Figure B shows a cluster of four stations positioned to give complete coverage within the 95% probability circle. The four Stations could not operate simultaneously without a substantial probability of interference, but the interference probability with contiguous clusters would be lower if the group of four was operated sequentially rather than simultaneously.

Figure C shows the increased distance separation of simultaneously operated Access-points when a group of 9 is operated sequentially and in step with contiguous like-type groups. Figure D is the same except the sequential group size is 16.

Each range circle that is double the radius of the interior circle corresponds to a change of about 10-12 dB/octave in average signal level.

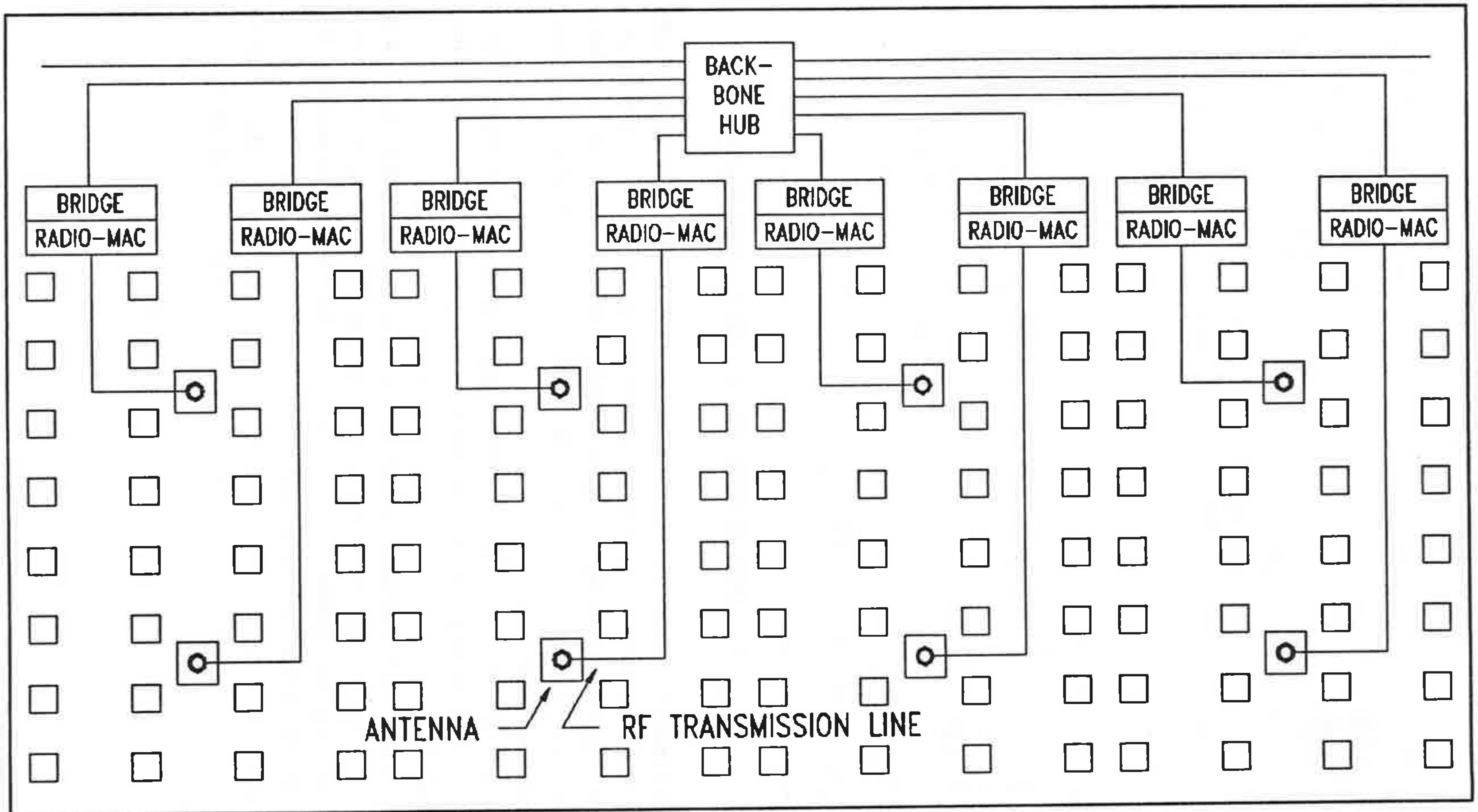

Figure 3 BRIDGE-PER-ACCESS-POINT WITH LAN BACKBONE

Topology is shown for a wireless access system based on using existing wired LAN for a backbone to interconnect Access-points. There is a backbone LAN, possibly either 802.3 CSMA/CD or 802.5 token ring, which might connect the Access-points directly ("daisy-chain") rather than looping through a common hub. Structured building wiring plans provide physical star wiring only which may be connected to appear electrically as a bus or ring. The bridge function at each Access-point is a relatively complex and costly equipment which is a major consideration in time required to adjust to the movement of stations from one Access-point to another.

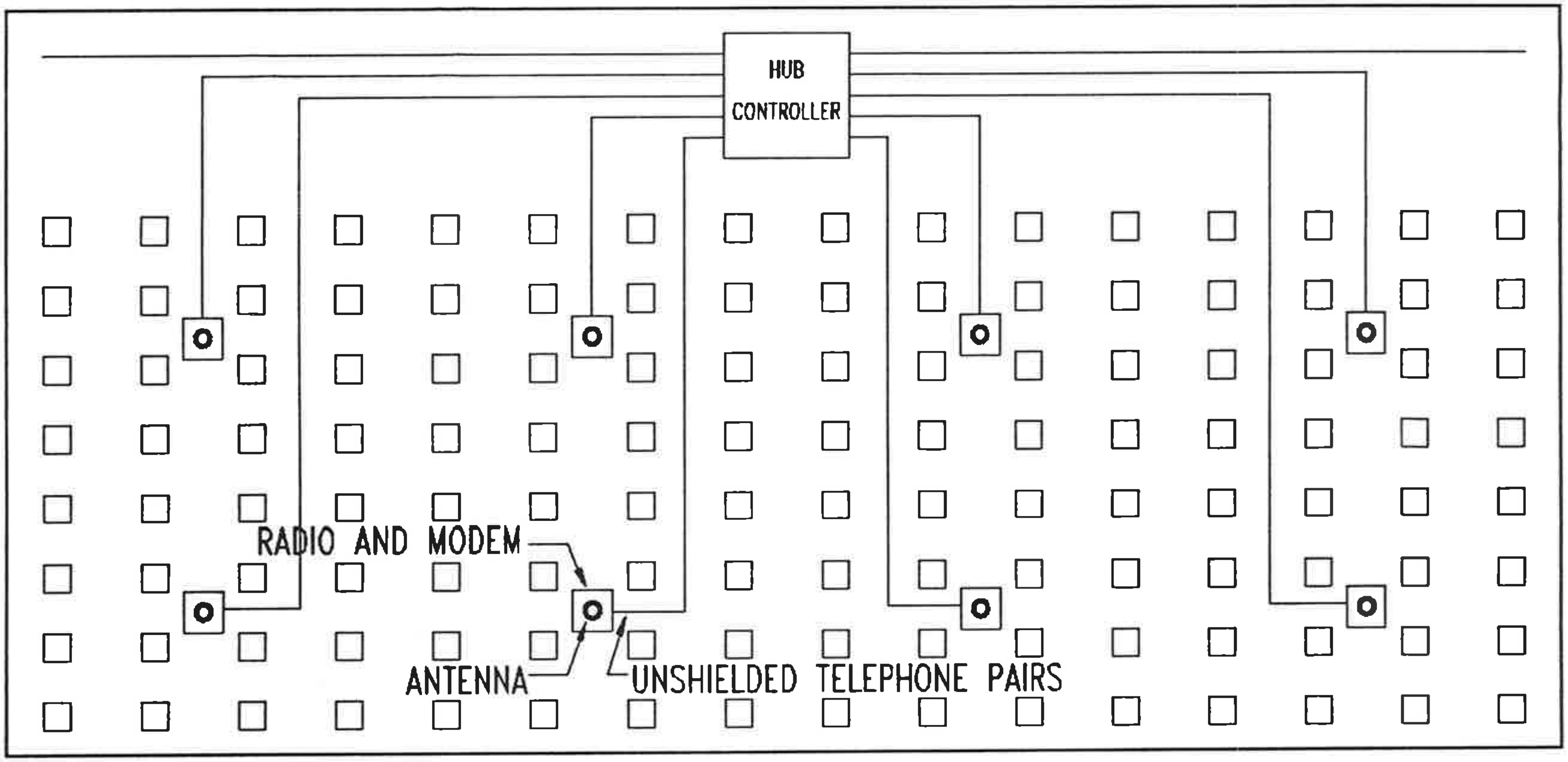

Figure 4 HUB CONTROLLER COMMON TO MANY ACCESS-POINTS

Topology is shown for a preferred wireless access system with a common Hub Controller for 8 Access-points each serving 16 Stations. The Access-point contains only the antenna, radio up-down converter and an intermediate frequency modem converting to/from baseband. The baseband analog signals pass over the telephone pairs where all further analog and digital processing takes place. The preponderance of electronics is centralized, and there are no transmission lines operated at radio frequency.