INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE CENTRATOM TRIESTE

The United Nations University

SMR.875-20

WORKSHOP ON TELEMATICS
2 - 20 October 1995

Background material for lectures
by

H. HASHEMI
Department of Electrical Engineering
Sharif University of Technology
P.O. Box 11365-9363
Tehran
IRAN

These are preliminary lecture notes, intended only for distribution to participants

THE BELL SYSTEM
TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING
ASPECTS OF ELECTRICAL COMMUNICATION

Volume 58 January 1979 Number 1

Advanced Mobile Phone Service: Introduction, Background, and Objectives

By W. R. YOUNG

(Manuscript received June 6, 1978)

This paper introduces a series of papers that describe in detail the Bell System's Advanced Mobile Phone Service (AMPS). It presents a brief history of mobile radio, highlighting the important events and legal decisions that preceded development of the AMPS system. The cellular system concept that has been embodied in AMPS makes large-scale mobile-radio service affordable to a sizable segment of the public. This concept calls for dividing transmission areas into "cells" to handle radio traffic, and, as traffic grows, subdividing those cells into smaller segments without increasing radio spectrum. This paper outlines AMPS objectives and sets the stage for more detailed articles on its evolution, its design and testing, and maintenance considerations.

I. INTRODUCTION

The potential for communicating with nonfixed points over the horizon without the use of wires was soon recognized following the invention of radio in the late 1800s and its development in the early 1900s. The first major use of this potential was to vessels at sea as an aid to navigation and safety. Since those early days, the use of mobile radio (as it is now called) has spread dramatically. Today it is used to communicate not only with ships at sea but with land vehicles, aircraft, and even with people using portable equipment.

The expanding need and concomitant growth have led to the development of the newest mobile system for common-carrier offering to the public, the Advanced Mobile Phone Service (AMPS),* the subject

* Known during developmental stages as High-Capacity Mobile Telecommunications System (HCMTS).

of this special issue of *The Bell System Technical Journal.* This system, in its mature configuration, will handle large quantities of mobile telephone traffic. High capacity will be achieved by dividing desired service areas into many small cells of radio coverage and, most important, by operating with the same radio spectrum utilized many times over within the service area.

This paper surveys the background and history of mobile radio, including governmental regulatory events, the development of systems used up to the present, and the emergence of new concepts and technology. The overview of AMPS introduces readers to the makeup, service objectives, and features of the new system.

II. BACKGROUND AND HISTORY

2.1 Early systems

In 1921 the Detroit police department made the earliest significant use of mobile radio in a vehicle.[1] That system operated at a frequency close to 2 MHz. The utility of this idea was so obvious that the channels in this low-frequency band were soon crowded.

New frequencies between 30 and 40 MHz were made available about 1940. A natural outgrowth of that development was the use of frequency modulation to improve reception in the presence of fading of the signal, electrical noise, and static. Opening the band encouraged a substantial buildup of police systems that started in the early 1940s and continues today.

Shortly thereafter, other users found a need for this form of communication. Private individuals, companies, and other public agencies purchased and operated their own mobile units and land (base) equipment. Over the years, the Federal Communications Commission (FCC) made available some 40 MHz of spectrum in bits and pieces between 30 and 500 MHz for various recognized and special uses. Today, approximately eight million licensed units enjoy this type of private service.* These systems are not generally connected directly to the telephone network.

In addition, the FCC has currently licensed over eight million citizens band radio units which are permitted to operate on 40 channels. An equal number of unlicensed units is also estimated to be operating on these channels. These figures graphically show that a great number of people want to communicate while on the move.

---

* In the early days, Bell System companies engineered, furnished, and maintained systems for private entities and public agencies such as police departments. This service was eliminated as a result of the 1956 consent decree: Final Judgment of January 24, 1956, in U.S.A. vs Western Electric et al.

2.2 Public correspondence systems

Immediately after World War II, the Bell System embarked on a program of supplying "public correspondence systems." The term means systems provided by a common carrier to permit communication among a variety of users that achieves large-scale economies by combining miscellaneous kinds of traffic into larger, more efficiently handled amounts. The FCC's official classification of this service is "Domestic Public Land Mobile Radio Service" (DPLMRS). (See Table I for a chronology of events in mobile radio history.)

The first of these public correspondence systems was inaugurated in 1946 to serve the city of St. Louis[2] with three channels near 150 MHz. The FCC had originally allocated six channels spaced 60 kHz apart, but the equipment was not sophisticated enough to prevent interference from adjacent channels being used in the same area. The St. Louis system was called an "urban" system.

In 1947, a "highway" system using frequencies in the 35- to 44-MHz band began operations along the highway between New York and Boston. These latter frequencies were thought to carry greater distances and, therefore, to be more useful in covering stretches of highway. However, these frequencies proved troublesome because of the skip-distance propagation phenomenon that carried unwanted conversations across the country. Today, the use of the 35- to 44-MHz band is declining.

Table I—History of mobile telecommunications related to common-carrier services

FCC Dockets		Service Offerings
	1946	First Bell System mobile service (150 MHz)
No. 8658, Bell System proposal for 40-MHz-bandwidth system	1947	Highway mobile service (35 MHz)
No. 8976, UHF TV, more detailed Bell System proposal for 40-MHz bandwidth system	1949	
	1956	First manual 450-MHz service
No. 11997, Bell System proposal for 75-MHz system at 800 MHz	1958	
	1964	First automatic 150-MHz service — MJ
	1969	First automatic 450-MHz service — MK
No. 18262, Allocation to common carriers		
— 75 MHz, tentative	1970	
— 40 MHz, firm	1974	
Open to "any" common carrier	1975	
Illinois Bell request for developmental authorization	1975	
Developmental authorization granted	1977	
	1978	AMPS Developmental System trial (850 MHz)

Both the urban and highway systems employed push-to-talk operation (somewhat unfamiliar to the ordinary telephone user) and were severely limited in the number of channels available. Nevertheless, more systems of both types were installed for cities and highways around the country. In many cases, the demand for service was such that the available channels could serve only a fraction of the demand for traffic and prospective customers had to be put on backlog lists.

Around 1955, the number of channels available at 150 MHz was expanded from 6 to 11 by the creation of new channels between old ones (i.e., channel spacing of 30 kHz). The year 1956 saw the addition of 12 channels near 450 MHz and the installation of the first system in this frequency range. All systems operated in the "manual" mode, with each call to or from a mobile unit handled by a special mobile operator. Mobile service still operates on a manual basis in some areas today.

In 1964, a new system, called the MJ, was developed and installed to improve efficiency, to reduce costs, and to achieve trunking advantage in cities having multiple channels. This system operated at 150 MHz, furnished automatic channel selection for each call, eliminated the need for push-to-talk, and allowed customers to do their own dialing. Most systems installed since 1964 are automatic, and many of the predecessor manual systems have been replaced.

In 1969 the automatic capability was extended to the 450-MHz channels with a system called the MK. The MJ and the MK were parts of the Improved Mobile Telephone System (IMTS),[3] the current standard for mobile service. In some respects, especially in convenience of dialing, the service given to IMTS customers is commensurate with that obtained with land-line telephones.

Present-day mobile telephone service requires a single land transmitter station positioned at a high elevation so that received signal levels at mobile units are substantially above the ambient noise throughout most of the desired coverage area. For each channel, the output power of the land transmitter is typically 200 or 250 watts, and transmitting antenna gain is sometimes used to raise the effective radiated power to 500 watts. Such a system ensures coverage as far as 20 or 25 miles from the transmitter site. Although the signal level on a channel may be poor beyond 25 miles, it is still high enough to interfere significantly with other mobile communications on the same frequency within 60 to 100 miles of the land transmitter. Consequently, two land transmitters spaced more closely than this should not use the same mobile telephone channel frequency. If land transmitters on the same frequency are farther apart, each can serve mobile units within about 20 miles with only minor interference, because any mobile unit is much closer to the land transmitter serving it than to any interfering transmitter.

From its inception to the present, mobile service has remained a

scarce luxury. Each month, mobile telephone customers typically pay 10 to 20 times as much for mobile service as for residential telephone service. Despite the cost, many telephone companies can cite long lists of "held orders"—unfilled requests for service—from people who want to become mobile subscribers. Market studies[4-7] have repeatedly uncovered a sizable demand at lower prices.

But even if the cost of mobile service could be reduced substantially, the primary factor that has hampered the spread of mobile service thus far has been the unavailability of spectrum. No new customers could be accommodated in many areas because only a few dozen channels are available for present-day service, and even these are fractured into several frequency bands and partitioned among different classes of service carriers.

Since 1949, common-carrier entities known as "Radio Common Carriers" (RCC), companies not providing public landline telephone service, have been given separate channels to furnish the same kind of mobile services as the wire-line common carriers (the Bell System and other telephone companies). With about the same number of channels available, they serve roughly the same number of customers. Table II shows the number of channels available for each type of carrier and the number of two-way mobile units served by each for the most recent year for which figures are available.

Compare the number (approximately 143,000) of RCC and wire-line common-carrier customers with the estimated 16 million or more private units not served by common carriers: the ratio is about 1:110.* There are many, both inside and outside the Bell System, who believe that this ratio reflects the number of available channels allocated to the different uses rather than the inherent demand for such services. The FCC has taken this into account in its most recent grant of 40 MHz of spectrum for use by common carriers.

2.3 Regulatory history

Since 1946, Bell System planners have been looking forward to the large-scale system they believed necessary to satisfy customer demands. Proposals for such a system were made from time to time, as described below. These generally were associated with FCC Dockets, as noted in the left-hand column of Table I.

In 1947, in connection with FCC Docket 8658, the Bell System asked for 12 more channels to use immediately in the same manner as the 6 already granted for urban service. Also requested was sufficient bandwidth for some 150 two-way channels from which large blocks of

* This is much smaller than the ratio of frequencies allocated (1:16) but is entirely consistent with the fact that the amount of traffic per mobile is much lower and channel loading is much higher in the private systems.

Table II—Channel allocations, number of mobile units, and number of systems

	Wireline Common Carriers			Radio Common Carriers	Total
	Bell	Independent	Total		
Number of two-way channels	23	23	23*	21†	54
MHz allocated	1.38	1.38	1.38	1.12	2.5
Number of mobile units (December 1977)	44,500	18,200	62,700	80,000‡	143,000 (approx.)
Number of systems (December 1977)	635	716	1,351	1,375	2,726

* Excludes 10 channels in the 35- to 40-MHz "highway" band, which are of limited and declining utility.

† Excludes the newest shared-with-TV channels in the 470- to 500-MHz band, since there has not been time for significant usage to build up.

‡ Projected forward from 1976 and earlier data.

channels could be assembled to achieve spectrum efficiency and capacity advantages. The planned 100-kHz spacing plus suitable guard bands (between mobile and land transmitters and between mobile and other services in adjacent bands) added up to approximately 40 MHz.

In 1949, a Bell System proposal representing a more mature plan for a broadband system was described in connection with FCC Docket 8976. This docket considered the disposition of UHF TV (470 to 890 MHz). The FCC decided at that time against providing a broadband mobile allocation in this band.

In 1958, the Bell System again made a broadband proposal, this time for a 75-MHz bandwidth (new estimated required spectrum) located at 800 MHz. This proposal was submitted as a response to an inquiry made by the FCC in its Docket 11997.

After considering the above proposal and the general pressure for more radio communications, in 1968 the FCC started Docket 18262, specifically addressed to the question of alleviating the large backlog of requests for frequencies for mobile use. Deliberating on requests for common-carrier service and for private-type service led the FCC to tentatively decide in 1970 to allocate 75 MHz for wire-line common-carrier use and 40 MHz to supplement private services. It proposed to do this by eliminating channels 70 through 83 in UHF TV and by using certain other pieces of spectrum from 806 to 947 MHz (a total of 115 MHz). The FCC invited industry to respond in 18 months with proposals for achieving communication objectives and demonstrating feasibility. In December 1971, the Bell System responded with a technical report which asserted feasibility by showing in considerable detail how a system might be composed.[8]

In 1974, the FCC made a firm allocation, different from the above: 40 MHz for wire-line common-carrier use and 30 MHz to supplement private services. The remainder of the 115 MHz was to be reserved pending further demonstrations of need. In doing this, the FCC strongly urged all suppliers to design their systems for greatest utility and spectrum efficiency.

Early in 1975, the FCC made some modifications in its 1974 decisions. One was to open the 40-MHz allocation for common-carrier service to "any qualified common carrier" rather than limit it to the wire-line carriers. In July 1975, the Illinois Bell Telephone Company filed a request to the FCC for authorization to install and test a developmental system in Chicago. This was granted in March 1977.

2.4 Emergence of key concepts

From our discussion thus far, it is obvious that the high-capacity system has been the result of planning and key concepts that have been emerging over a long period of time. Perhaps the first concept to be appreciated as necessary to an efficient, large-capacity operation was trunking, so much so that it was part of the proposal to the FCC in 1947. Trunking, as used here, is the ability to combine several channels into a single group so that a mobile can be connected to any unused channel in the group for either an incoming or outgoing call. This arrangement reduces blocking probability and greatly increases traffic-carrying efficiency relative to the situation in which a mobile unit can utilize only one fixed channel.*

One problem that bothered early planners was how to achieve full trunking advantage without requiring each mobile unit to be able to tune to every one of the channels in use for this service throughout the country. In those days, each new operating frequency required two quartz crystals and a position on the channel selector switch. The solution came when it became technologically feasible to construct a low-cost frequency synthesizer that could be set on any of a large number of frequencies but required only a small number of quartz crystals. While the basic idea is quite old, the circuit was made practical and economical only in the early 1970s. It is now taken for granted in ongoing planning.

The *cellular* concept and the realization that small cells with spectrum re-use could increase traffic capacity substantially seem to have materialized from nowhere, although both were verbalized in 1947 by D. H. Ring of Bell Laboratories in unpublished work. According to the

* The IMTS systems employ trunking to advantage, but the small number of channels in use in a given system (typically less than the 12 that could be assigned) limits trunking efficiency.

cellular concept, a desired service area is divided into regions called cells, each with its own land radio equipment for transmission to and from mobile units within the cell. It was further recognized that if the available channels were distributed among smaller cells the traffic capacity would be greater. Thus a system needing a relatively small capacity could use large cells, and, as necessary to achieve larger capacity, these cells could be divided into smaller ones. Each channel frequency can then be used for many independent conversations in many cells which are spaced far enough from each other to avoid undue interference.

From 1947 on, the teams planning the eventual system had faith that the means for administering and connecting to many small cells would evolve by the time they were needed. Those means did, in fact, become a reality with the advent of electronic switching technology.

Locating and *handoff* are concepts that come directly from the use of small cells. The act of transferring from one channel to another is called handoff. "Locating" is a process for determining whether it would be better from the point of view of signal quality and potential interference to transfer an active connection with a mobile unit to another land transmit/receive equipment, or perhaps to another land site.* The process entails sampling the signal from the mobile unit to determine if handoff from one voice channel to another is required. Since a mobile unit will sometimes move beyond the borders of one cell into another, it will be desirable to transfer the connection to an appropriate new cell.

The system, as presently planned, uses omnidirectional antennas when the cells are large. When smaller cells are created, directional antennas are used which divide each cell into three sectors, each served by an appropriate directional antenna at the cell site. This concept was introduced many years ago.† This advantageous arrangement reduces the amount of co-channel interference from surrounding cells and increases system capacity. It is covered further in Ref. 9.

The plan for increasing traffic capacity, as required, from a sparse system to a mature system in a given metropolitan service area, assumes the division of the large cells used at first into small cells as needed. The best method for achieving this is a growth plan developed in recent years (see Ref. 9).

III. OVERVIEW

This section gives an overview of the AMPS system, covering the objectives, the basic system, services and features, and additional problems and considerations.

---

* The prime purpose of this process is not to determine the geographic location of the mobile unit, although the geographical location is a statistical factor in performance.

† Described in Ref. 8.

3.1 Objectives

The major AMPS system objectives are discussed in the following paragraphs.

(*i*) *Large subscriber capacity:* The capability of serving a large amount of traffic to many thousands of mobile users within a local service area, such as a greater metropolitan area, within a fixed allocation of several hundred channels is essential to AMPS.

(*ii*) *Efficient use of spectrum:** The scarcity of radio spectrum as a public resource demands that it will be used responsibly. AMPS will use it efficiently, for unless this is achieved, AMPS would lack the ability to take care of the large anticipated traffic within the allotted band.

(*iii*) *Nationwide compatibility:* The FCC strongly urges nationwide compatibility. The objective means that mobile systems everywhere should provide the same basic service with the same standards of operation to be sure that a mobile station based in one place will achieve satisfactory service elsewhere.

(*iv*) *Widespread availability:* Studies of existing services show that it is important to many users to be able to roam far from their normal home system and still receive service. Neither this characteristic nor nationwide compatibility necessarily implies universal coverage. Wide-area coverage will be achieved gradually as metropolitan systems extend their coverage into surrounding suburbs, and finally along the principal road and rail routes between metropolitan centers.

(*v*) *Adaptability to traffic density:* Since the density will differ from one point in an area to another in a city and more remote points, and since all of this will change with time, an AMPS objective is to be adaptable to these variable needs.

(*vi*) *Service to vehicles and to portables:* While AMPS is conceived primarily for use with vehicles, an important objective is to make it compatible with portables (hand-carried). This should be possible with little or no compromise in the design of the land-based network.

(*vii*) *Regular telephone service and special services, including "dispatch":* In addition to regular telephone service, AMPS should provide specialized services, such as dispatch or fleet operation, and special features, such as abbreviated dialing.

(*viii*) *"Telephone" quality of service:* As for quality of service, the capability objective is essentially the same quality as ordinary

---

* A meaningful measure of spectrum efficiency is the number of simultaneous voice-communication paths that can be created per megahertz of spectrum and per square mile of area. This measure is useful where mobile terminals are statistically scattered throughout a service area.

nonradio telephone service. Since the types of impairments encountered are not always the same, it is sometimes difficult to ensure achieving identical quality. The goal is that the audio quality—faithful reproduction of voice and freedom from excessive noise and distortion—will not differ in overall effect as perceived by the user. It also means that service quality as measured by occasional blocking of the paths from customer to central office will not be noticeably greater than that encountered in the land network. This will be a very large improvement over current radio service, in which the pressure to accommodate many customers results in channel loading which frequently causes the probability of blocking to exceed 50 percent.

(*ix*) *Affordability:* A goal is to make the service affordable by a substantial portion of the public and of businesses. Cost economies due to large production runs will tend to make this possible.

3.2 Basic system

Figure 1 shows the basic structure of the system as presently planned. The service area to be covered is divided into an appropriate number of cells. Each cell site has radio equipment and associated controls that can effect the connection to any mobile unit located in the cell. The cell sites are interconnected to and controlled by a central Mobile Telecommunications Switching Office (MTSO). The MTSO is basically a telephone switching office with substantial capabilities for software control. It connects to the telephone network and also provides the means to perform maintenance and testing and to record call information for billing purposes.*

All of the above make up the land-based part of the AMPS system. The mobile units complete the system.

The frequency layout (channel assignments) plan, the plan of operation for the system, and the way objectives cited earlier will be achieved are described in Ref. 9.

3.3 Services and features

The basic service is a telephone in a vehicle and is analagous to the individual telephone in the nationwide telephone network.

Beyond this, the intention is to offer mobile users features ordinarily available to telephone users, with emphasis on those of particular value in the mobile environment. One feature not generally available

---

* In this overview, the MTSO is portrayed as a compact monolithic entity. In future practice, however, there may be multiple MTSOs, as required for achieving greatest economy and sufficient capacity.

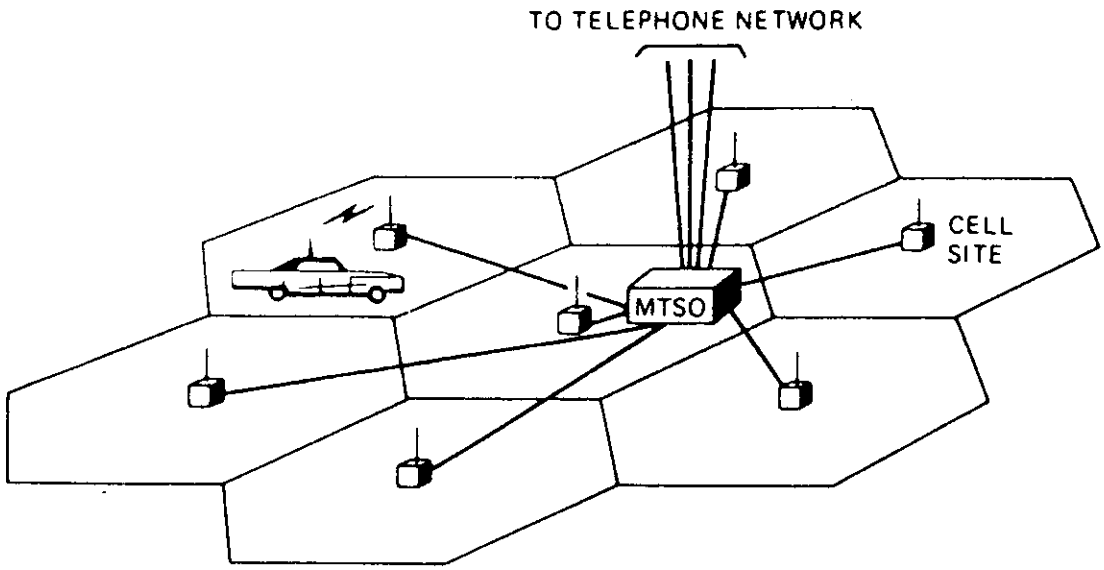


Fig. 1—The components and layout of the AMPS system.

in non-mobile phones, *pre-origination dialing*, will be included. This allows a customer to enter and store the called destination number before going off-hook. Then, when the user wishes to place the call, he begins the connection process by going off hook; the system uses this stored number to complete the connection. If the called line is busy or doesn't answer, the user may try again later without having to enter the same number again.

Eventually, the following vertical services, *Customer Calling Services,* furnished by ESS offices[10] may be made available for mobile users:

(*i*) "Three-Way Calling" permits a mobile user whose phone is already connected to another phone to originate a call to a third party, to switch back and forth between the connections, to bridge both connections as desired, or to connect the two other parties for continued conversation and then disconnect himself. With this feature, the mobile user can, for example, transfer a connection to another party.

(*ii*) "Call Waiting" furnishes a signal to alert the mobile user to an incoming call while a conversation is already in progress. By making use of the three-way calling feature, the customer will be able to transfer, accept the new call, and hold or terminate the former connection in progress.

(*iii*) "Speed Calling" permits a customer to originate a call to any of a few frequently called numbers by pushing one or two buttons. The connection is completed by the ESS in accordance with information stored there. This feature should be especially useful in a vehicle where a user cannot conveniently consult a directory or written notes. This feature will be implemented in the first working system.

In addition, it is expected that more features will be made available which can be implemented within the design of the mobile equipment. An example is "*Repertory Dialing*," which is similar to "Speed Calling" except that the mobile equipment stores the numbers and completes the calls. This feature is more useful than speed calling for vehicles that roam from system to system and, therefore, need to carry their own repertory.

3.4 Additional considerations

Other plans for mobile communications include numbering and dialing, the provision for roaming from system to system, the provision of operator service, and tariffs and billing.

3.4.1 Numbering and dialing

Each mobile unit is assigned a 10-digit number (including area code). The mobile user will dial seven or ten digits with a 0 or 1 prefix, where applicable, as if calling from a fixed telephone. The adopted numbering plan places no requirements on the overall nationwide numbering plan; for example, no special office or area codes need be set aside to separate mobile traffic from other telephone traffic.

3.4.2 Roaming

A strong need for serving vehicles that roam has been identified. This capability is needed not just within a greater metropolitan service area, but to other service areas and along the highways between. The roaming capability will not be demonstrated in the Chicago trial,* but a method of operation for systems beyond the trial has been planned. Wherever there is a system to serve it, a mobile unit will be able to obtain completely automatic service.

However, a call from a land telephone to a mobile unit which has roamed to another metropolitan area presents additional problems. While it would be logically possible for the system to determine automatically where the mobile unit is, and to connect it to the land party, there are two reasons for not doing so. First, the land customer will expect only a local charge if the mobile unit's number is a local one, and the mobile customer may not wish to pay the toll difference. Second, the mobile user may not want to have his whereabouts divulged through this system, automatically, without his permission. To respect the customer's wishes in this regard, the system will complete the connection only if the extra charge is agreed to, and only where it is possible without unauthorized disclosure of the service area to which the mobile unit has roamed.

---

* Identified in Section 3.5.

3.4.3 Operator services

Standard operator services will be available to mobile users. No special operators or operator services (except possibly for handling the roaming situation) will be required for the AMPS system.

3.4.4 Tariffs and billing

The MTSO will record connect and disconnect times, location information, and call-destination information as required for billing. The recorded information will be tailored to the needs of tariff and charging algorithms, when these have been determined.

3.5 System tests and trial

Tests to provide information for system planning, establishing feasibility, and implementing of AMPS have been conducted relevant to different aspects of the AMPS system. Most of these were directed at learning about radio propagation, radio noise and interference, antenna characteristics and performance, etc.

Currently there are two major "tests" of the AMPS system: (1) The Cellular Test Bed (CTB) in Newark, N.J., and (2) the Developmental System in Chicago, Ill. Since these are described fully in Refs. 11 and 12, respectively, in this issue, the discussion here is kept brief. Suffice it to say that the former is a system laid out geographically to simulate a mature cellular system and permit measurements of coverage and interference in an actual urban layout. It does not simulate the whole service involving mobile customers, but is a "laboratory in the field." The Developmental System is an initial installation of a system implemented to serve mobile users, and will demonstrate the service itself as well as its implementation. But since the Developmental System employs relatively few large cells, it is not intended to demonstrate operation of a fully mature, small-cell layout. These two major endeavors complement each other and, taken together, provide a demonstration of all the major features of AMPS.

As explained later, the trial of the Developmental System has two phases. The first, which started in July 1978, is called an "Equipment Test," has about 100 mobile units, and is intended to "shake-down" the system and demonstrate that the system operates satisfactorily. The second phase, following the Equipment Test, is called a "Service Test," involving approximately 2100 users, and will demonstrate the service aspects of the system.

IV. SUMMARY

This first paper of the series has provided a general introduction to and overview of AMPS. The papers that follow describe more com-

pletely the cellular concept, control architecture, voice and data transmission aspects, the Cellular Test Bed, and the Developmental System. Other papers describe the Mobile Telecommunications Switching Office, the subscriber set used in the equipment-test phase of the developmental system operation, the mobile telephone control unit used in the service test, the hardware used at cell sites, and laboratory test systems that were devised to obtain operational data during system tests.

REFERENCES

1. W. C. Jakes, ed., *Microwave Mobile Communications*, New York: John Wiley, 1974.
2. "Telephone Service for St. Louis Vehicles," Bell Laboratories Record, *24*, No. 7 (July 1946), pp. 267–268.
3. H. W. Nylund and R. M. Swanson, "Improved Mobile Dial Telephone System," IEEE Trans. Veh. Commun., *1* 1963, pp. 32–36.
4. "Communications Tomorrow: Who? What? Which? Where? Why? and How Much?" Communications; Part I (October 1972), pp. 42–45; Part II (November 1972), pp. 20–23.
5. "World Mobile Telecommunications," Cambridge, Mass.: Arthur D. Little, 1974.
6. "A New Era in Mobile Communications," Raymond Kraus, Consulting Communications Engineers, Inc., 1976.
7. "The U.S. Mobile Equipment Market," Report No. 556, May 1978, Frost and Sullivan, Inc.
8. Bell Laboratories, "High-Capacity Mobile Telephone System—Technical Report," (December 1971), submitted to the FCC.
9. V. H. MacDonald, "AMPS: The Cellular Concept," B.S.T.J., this issue, pp. 15–41.
10. Bell Laboratories Record, *46*, No. 6 (June 1965), page 210.
11. G. C. DiPiazza, A. Plitkins, and G. I. Zysman, "AMPS: The Cellular Test Bed," B.S.T.J., this issue, pp. 215–248.
12. D. L. Huff, "AMPS: The Developmental System," B.S.T.J., this issue, pp. 249–269.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 58, No. 1, January 1979
Printed in U.S.A.

Advanced Mobile Phone Service:

The Cellular Concept

By V. H. MAC DONALD

(Manuscript received July 17, 1978)

This paper shows how a cellular system operating within a limited block of frequency spectrum can meet the objectives of a large-scale mobile-telephone service designed with attention to cost restraint. It explores the key elements of the cellular concept—frequency reuse and cell splitting—and describes certain mathematical properties of hexagonal cellular geometry. A description of the basic structure and features of AMPS *shows how the cellular concept can be put into practice.*

I. INTRODUCTION

The preceding paper in this issue[1] noted that Bell System planners were already looking ahead to a more economical and widespread form of mobile-telephone service when early mobile telephone systems were being installed in the 1940s. Since then, system designers have recognized that a substantial block of radio-frequency spectrum, equivalent to hundreds of voice channels, is a prerequisite for a large-scale mobile service. This spectrum was provided by the FCC's reallocation of a portion of the former UHF television band for mobile service in Docket 18262. This paper cites the system objectives adopted over the years, explores the cellular concept which evolved in response to these objectives, and describes many aspects of a practical embodiment of the cellular concept—the Advanced Mobile Phone Service (AMPS) system.

II. OBJECTIVES FOR LARGE-SCALE MOBILE-TELEPHONE SERVICE

Over the years, system designers have set various objectives for large-scale mobile-telephone service, based on the interests of the

public, mobile-telephone customers, and mobile-telephone operating companies. The first paper in this issue[1] cited the following basic objectives:

(*i*) Large subscriber capacity.
(*ii*) Efficient use of spectrum.
(*iii*) Nationwide compatibility.
(*iv*) Widespread availability.
(*v*) Adaptability to traffic density.
(*vi*) Service to vehicles and portables.
(*vii*) Regular telephone service and special services, including "dispatch."
(*viii*) "Telephone" quality of service.
(*ix*) Affordability.

Various systems might be devised to satisfy all the above objectives, except for the first two. The system must be capable of growing to serve many thousands of subscribers within a local service area, such as the environs of a single city, yet the provision of service must not be contingent on the continual enlargement of the allocated spectrum. The need to operate and grow indefinitely within an allocation of hundreds of channels has been the primary driving force behind the evolution of the cellular concept.

III. BASIC ELEMENTS OF THE CELLULAR CONCEPT

The two phrases *frequency reuse* and *cell splitting* summarize the essential features of the cellular concept.

3.1 Frequency reuse

Frequency reuse refers to the use of radio channels on the same carrier frequency to cover different areas which are separated from one another by sufficient distances so that co-channel interference is not objectionable. Frequency reuse is employed not only in present-day mobile-telephone service but also in entertainment broadcasting and most other radio services.

The idea of employing frequency reuse in mobile-telephone service on a shrunken geographical scale hints at the cellular concept. Instead of covering an entire local area from one land transmitter site with high power at a high elevation, the service provider can distribute transmitters of moderate power throughout the coverage area. Each site then primarily covers some nearby subarea, or zone, or "cell." A cell thus signifies the area in which a particular transmitter site is the site most likely to serve mobile-telephone calls. Figure 1 is a sketch of a cellular map or "layout." In principle, the spacing of transmitter sites does not need to be regular, and the cells need not have any particular shape. Cells labeled with different letters must be served by distinct sets of channel frequencies to avoid interference problems. A cell

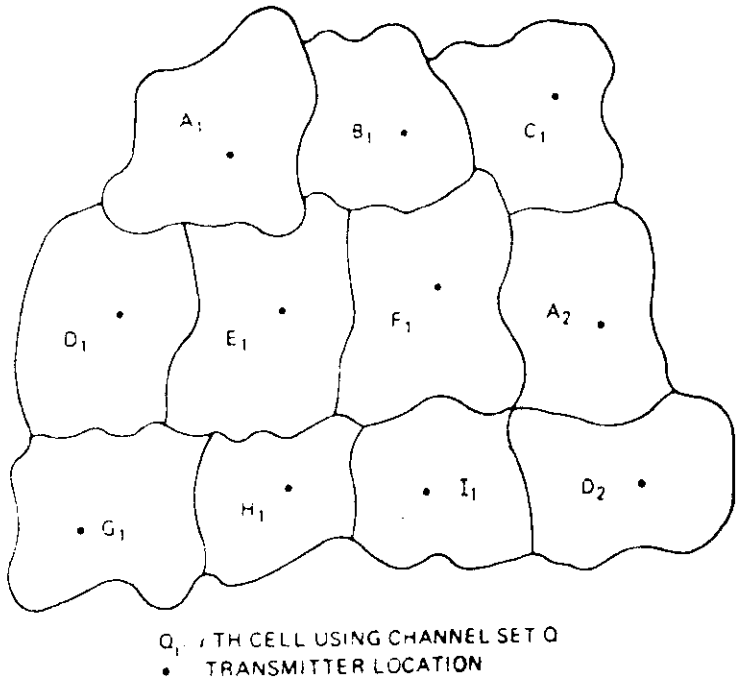


Fig. 1—Cellular layout illustrating frequency reuse.

therefore has the additional significance that it is the area in which a particular channel set is the most likely set to be used for mobile-telephone calls. Cells sufficiently far apart, such as those labeled A_1 and A_2, may use the same channel set.

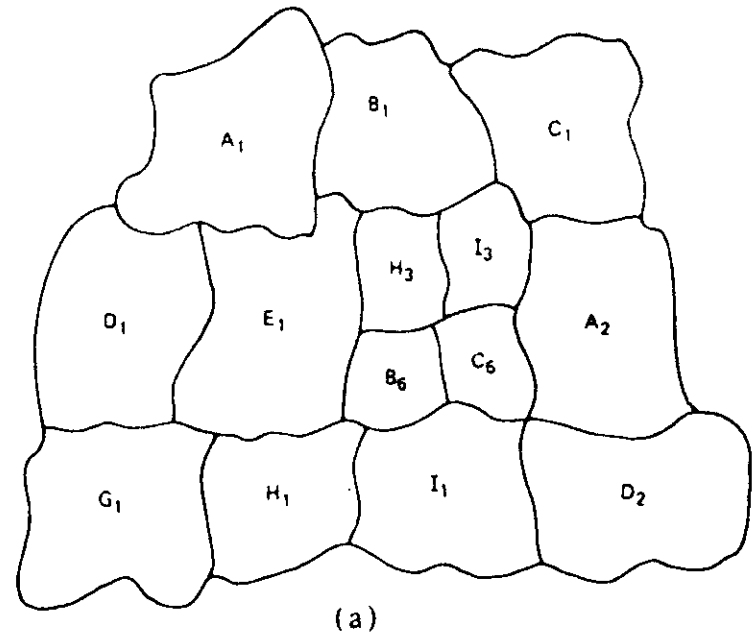
Through frequency reuse, a cellular mobile-telephone system in one coverage area can handle a number of simultaneous calls greatly exceeding the total number of allocated channel frequencies. The multiplier by which the system capacity in simultaneous calls exceeds the number of allocated channels depends on several factors, particularly on the total number of cells.

3.2 Cell splitting

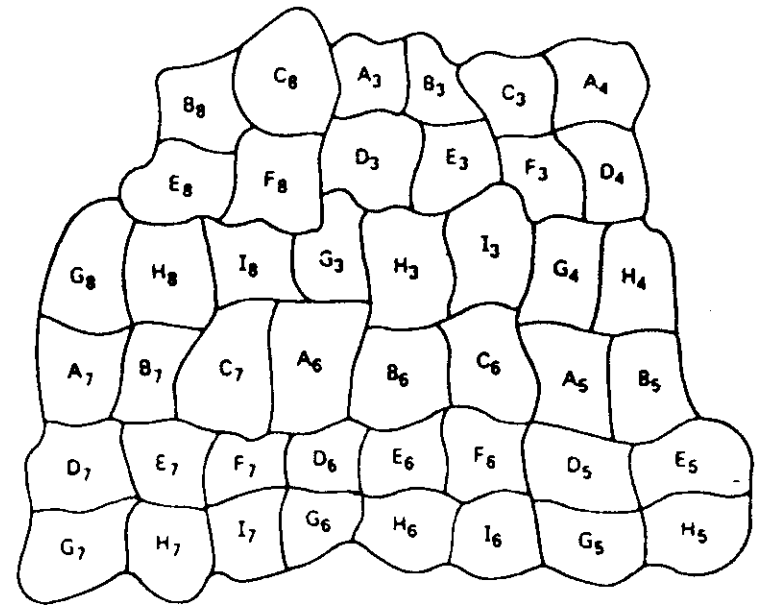
If the total allocation of C channels is partitioned into N sets, then each set will contain nominally $S = C/N$ channels. If one channel set is used in each cell, eventually the telephone traffic demand in some cell will reach the capacity of that cell's S channels. Further growth in traffic within the cell will require a revision of cell boundaries so that the area formerly regarded as a single cell can now contain several cells and utilize all these cells' channel complements. The process called "cell splitting" fills this need.

Figure 2a illustrates an early stage of the cell-splitting process, in which the cell originally designated F_1 (in Fig. 1) has reached capacity. The area previously treated as cell F_1 now contains cells H_3, I_3, B_6, and C_6. If the demand in the area continues to grow, other larger cells will be split, and eventually, as in Fig. 2b, the entire region will be converted into smaller cells.

In practice, splitting a given cell may be less abrupt than our



(a)



Q_i: iTH CELL USING CHANNEL SET Q

(b)

Fig. 2—Cellular layout illustrating cell splitting. (a) Early stage. (b) Later stage.

illustration implies. It is often sufficient initially to superimpose just one or two smaller cells onto a larger cell, so that the larger and smaller cells jointly serve the traffic within the area spanned by the smaller cell(s). The larger cell disappears at a later time, when all its territory becomes covered by smaller cells. We discuss this aspect of system growth in Section 7.3. In Figs. 1 and 2, for illustration the total allocation has been partitioned into nine distinct channel sets, labeled *A* through *I*. The figures show a progression from an initial stage (Fig. 1), in which each allocated channel is available once within the region spanned by cells A_1 through I_1, to a later stage (Fig. 2b), in which each channel is available in four different cells within that same region.

Successive stages of cell splitting would further multiply the number of "voicepaths," i.e., the total number of simultaneous mobile-telephone calls possible within the same region. By decreasing the area of each cell, cell splitting allows the system to adjust to a growing spatial traffic demand density (simultaneous calls per square mile) without any increase in the spectrum allocation.

The techniques of frequency reuse and cell splitting permit a cellular system to meet the important objectives of serving a very large number of customers in a single coverage area while using a relatively small spectrum allocation. Cell splitting also helps to meet the objective of matching the spatial density of available channels to the spatial density of demand for channels, since lower-demand areas can be served by larger cells at the same time that higher-demand areas are served by smaller cells.

IV. PROPERTIES OF CELLULAR GEOMETRY

The main purpose of defining cells in a mobile-telephone system is to delineate areas in which either specific channels or a specific cell site will be used at least preferentially, if not exclusively. A reasonable degree of geographical confinement of channel usage is necessary to prevent co-channel interference problems. Having defined a desired cellular pattern in concept, system planners achieve that pattern in the field through proper positioning of land transmitter sites, proper design of the azimuthal gain pattern of the sites' antennas, and proper selection during every call of a suitable site to serve the call.

The irregular land transmitter spacing and amorphous cell shapes shown in Figs. 1 and 2 might be acceptable in a system where the initial system configuration, including the selection of transmitter sites and the assignment of channels to cells, could be frozen indefinitely. In practice, however, the absence of an orderly geometrical structure in a cellular pattern would make adaptation to traffic growth more cumbersome than necessary. Inefficient use of spectrum and uneconomical deployment of equipment would be likely outcomes. A great deal of improvisation and custom engineering of radio, transmission, switching, and control facilities would be required repeatedly in the course of system growth.

Early in the evolution of the cellular concept, system designers recognized that visualizing all cells as having the same shape helps to systematize the design and layout of cellular systems. A cell was viewed as the coverage area of a particular land site. If, as with present-day mobile service, omnidirectional transmitting antennas were used, then each site's coverage area—bounded by a contour of constant signal level—would be roughly circular. Although propagation considerations recommend the circle as a cell shape, the circle is impractical

for design purposes, because an array of circular cells produces ambiguous areas which are contained either in no cell or in multiple cells. On the other hand, any regular polygon approximates the shape of a circle and three types, the equilateral triangle, the square, and the regular hexagon, can cover a plane with no gaps or overlaps (Fig. 3). A cellular system could be designed with square or equilateral triangular cells, but, for economic reasons, Bell Laboratories system designers adopted the regular hexagonal shape several years ago.

To understand the economic motivation for choosing the hexagon, let us focus our attention on the "worst-case" points in a cellular grid—the points farthest from the nearest land site. Assume a land site located at the center of each cell, the center being the unique point equidistant from the vertices. The vertices are in fact the worst-case points, since they lie at the greatest distance from the nearest land site. Restricting the distance between the cell center and any vertex to a certain maximum value helps to assure satisfactory transmission quality at the worst-case points. If an equilateral triangle, a square, and a regular hexagon all have the same center-to-vertex distance, the hexagon has a substantially larger area. Consequently, to serve a given total coverage area, a hexagonal layout requires fewer cells, hence fewer transmitter sites. A system based on hexagonal cells therefore costs less than one with triangular or square cells, all other factors being equal.

With our present understanding of cellular systems, we recognize that, because of propagation vagaries, it is not possible to precisely define a coverage area for a given cell site in the sense that the site never serves mobile units outside the area and always serves mobile units within the area. Nevertheless, the concept of a cell remains valid in the context of an area in which a certain land site is more likely to serve mobile-telephone calls than any other site.

A familiarity with some of the basic properties of hexagonal cellular geometry will give the reader additional perspective on subsequently

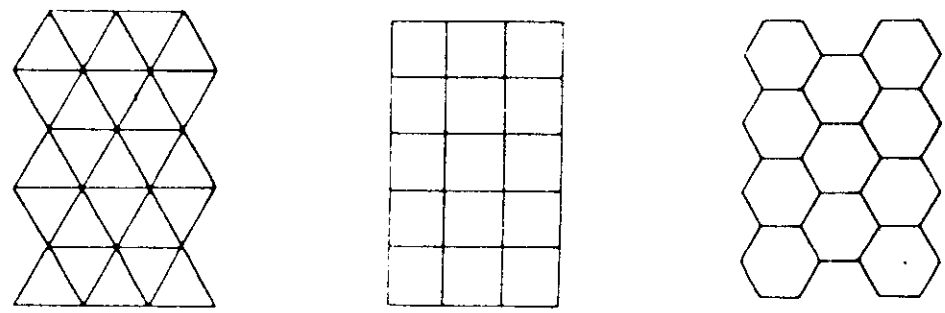


Fig. 3—Regular polygons as cells. (a) Equilateral triangles. (b) Squares. (c) Regular hexagons.

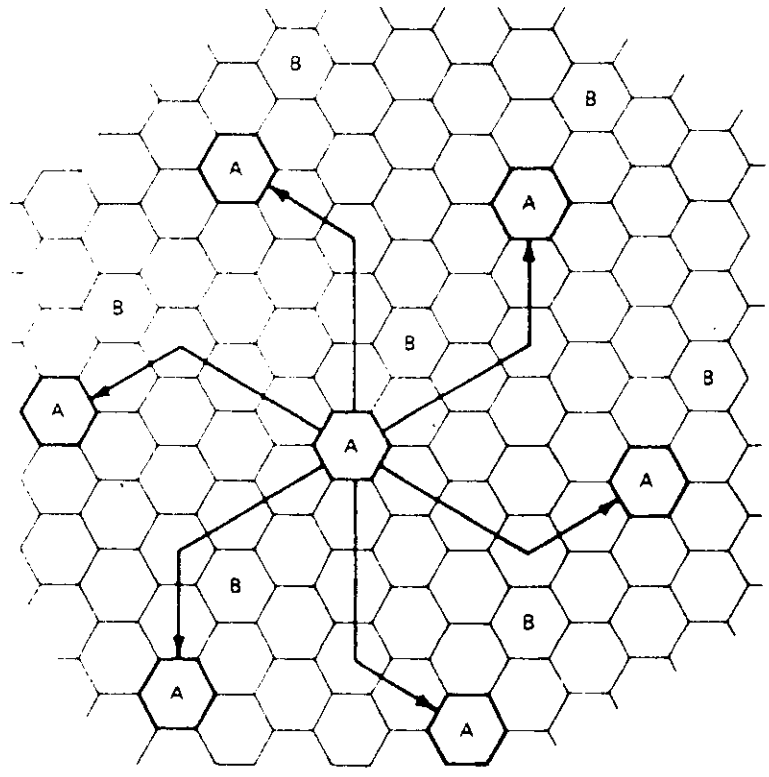
discussed details of radio-channel assignment in a cellular system. We now explain how cells using the same channel set are oriented with respect to one another and how cellular patterns and certain basic geometrical parameters are related to one another.

To lay out a cellular system in the sense of determining which channel set should be assigned to each cell, we begin with two integers i and j ($i \geq j$), called "shift parameters," which are predetermined in some manner. From the cellular pattern of Fig. 4, note that six "chains" of hexagons emanate from each hexagon, extending in different directions. Starting with any cell as a reference, we find the nearest "co-channel" cells, that is, those cells that should use the same channel set, as follows:

> Move i cells along any chain of hexagons; turn counter-clockwise 60 degrees; move j cells along the chain that lies on this new heading.

The jth cell and the reference cell are co-channel cells. Now return to the reference cell and set forth along a different chain of hexagons using the same procedure.

Figure 4 illustrates the use of these directions for an example in which $i = 3$ and $j = 2$. A cell near the center of the figure is taken as a reference and labeled A. As each co-channel cell is located, it is also



SHIFT PARAMETERS, $i = 3, j = 2$

Fig. 4—Illustration of the determination of co-channel cells.

labeled A. To continue the cellular layout, one could choose another label, such as B, for a cell close to the reference cell and find this cell's nearest co-channel cells. However, once the position of all the cells labeled A is determined, it is not necessary to work through the procedure described above for subsequent labels. The pattern of cell labels built up around the reference A cell is simply replicated around all the other A cells by translation without rotation.

Co-channel cells could also be located by moving j cells before turning and i cells afterwards, rather than vice versa, or by turning 60 degrees clockwise instead of counterclockwise. There are four different ways of describing the procedure, and two different configurations can result. Each configuration is just the reflection of the other across an appropriate axis.

When a sufficient number of different labels has been used, all cells will be labeled, and the layout will be complete. The cells form natural blocks or clusters around the reference cell in the center and around each of its co-channel cells. The exact shape of a valid cluster is not unique; all that is required is that it contain exactly one cell with each label. The number of cells per cluster is a parameter of major interest, since in a practical system this number determines how many different channel sets must be formed out of the total allocated spectrum. The number of cells per cluster, N, turns out to be

$$N = i^2 + ij + j^2. \tag{1}$$

(The appendix to this paper derives this result and presents additional information on hexagonal cellular geometry.) The fact that i and j must be integers means that only certain values of the number of cells per cluster are geometrically realizable.

The ratio of D, the distance between the centers of nearest neighboring co-channel cells, to R, the cell radius, is sometimes called the "co-channel reuse ratio." This ratio is related to the number of cells per cluster, N, as follows:

$$D/R = \sqrt{3N}. \tag{2}$$

In a practical system, the choice of the number of cells per cluster is governed by co-channel-interference considerations. As the number of cells per cluster increases, the relative separation between co-channel cells obviously increases, and consequently poor signal-to-interference conditions become progressively less probable. Section 6.4 discusses a method for choosing the number of cells per cluster.

V. AMPS: A PRACTICAL REALIZATION OF THE CELLULAR CONCEPT

This section describes the physical structure of the AMPS system and provides a glimpse of its basic control algorithms to show how cellular operation can be effected in a working system.

The implementation of the cellular concept in a practical system requires the construction of an essentially regular array of land transmitter-receiver stations, called "cell sites" in AMPS. The design abstraction of an array of cells is embodied in the physical reality of the cell-site array. The dots in Fig. 5a symbolize an idealized AMPS cell-site array, consisting of a lattice of regularly spaced cell sites. For this idealized array of cell sites, an accompanying pattern of regular hexagonal cells can be visualized in at least two different ways: (*i*) cells whose centers fall on cell sites, "center-excited" cells (Fig. 5b), or (*ii*) cells half of whose vertices fall on cell sites, "corner-excited" cells (Fig. 5c).

Section 6.1 acknowledges the practical reality that it is seldom possible to position a cell site exactly at its geometrically ideal location and discusses the degree to which the actual location may deviate from the ideal.

Center-excited cells exemplify the previous practical definition of a cell as the area in which one particular cell site is more likely to be used on mobile-telephone calls than any other site. On any single call, neither the mobile unit's nor the system's actions would clearly delineate any cell boundaries, but a protracted study of system behavior would reveal the presence of center-excited cells satisfying the above pragmatic definition. Because of random propagation effects, any real cell only approximates the ideal hexagonal shape but, for purposes of design and discussion, it is appropriate to visualize cells as regular hexagons. The practical meaning of a corner-excited cell will be explained in conjunction with the ensuing discussion of directional cell sites.

5.1 Omnidirectional and directional cell-sites

The AMPS plan envisions that, at the inception of the system in any locality, the cell sites will use transmitter and receiver antennas whose patterns are omnidirectional in the horizontal plane.[3] The use of omnidirectional antennas has traditionally been depicted by the cen-

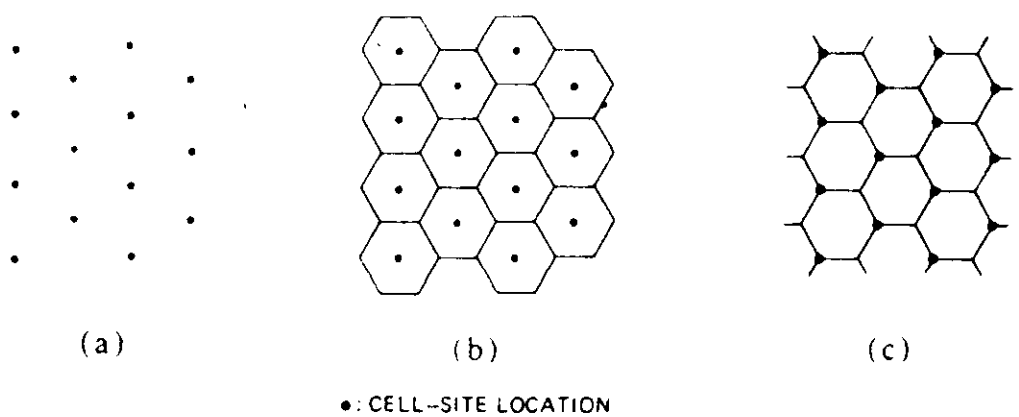


Fig. 5—Cellular geometry with and without cells. (a) Cell-site lattice. (b) "Center-excited" cells. (c) "Corner-excited" cells.

ter-excited cell pattern of Fig. 5b. The phrase "omnidirectional cell site" refers to a site equipped with omnidirectional voice-channel antennas.

In mature systems, cell sites will have three faces, that is, each voice channel in a cell site will be transmitted and received over one of three 120-degree sector antennas, rather than over an omnidirectional antenna. The antennas will be oriented as shown in Fig. 6, so that extensions of the edges of the antennas' front lobes form the sides of hexagonal cells as in Fig. 5c. These are the "corner-excited" cells that have customarily been employed to suggest the tri-directional coverage of AMPS cell sites in mature systems.

Cell sites are very expensive investments. The initial cost of a site, before installation of any voice-channel transceivers, is much greater than the incremental cost of each subsequently installed voice channel. At the inception of a system, the number of sites is governed strictly by the need to span the desired coverage area. At this stage, omnidirectional sites are used because the initial cost of an omnidirectional site is lower than that of a directional site. In mature systems, however, the potential, explained below, for cutting cost by reducing the total

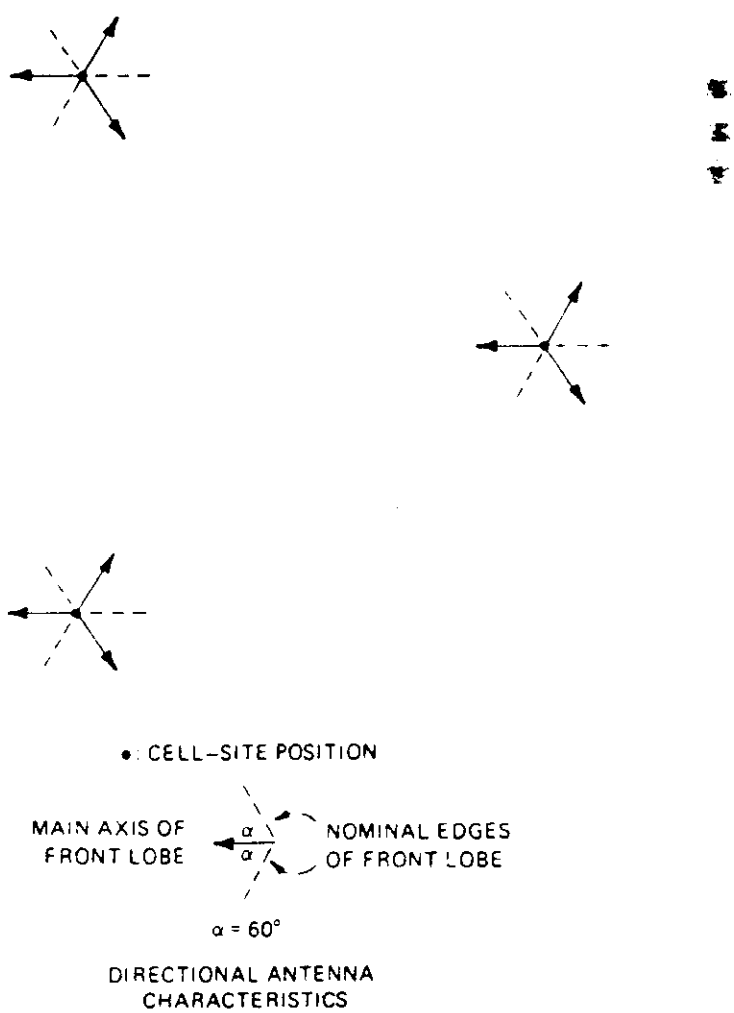


Fig. 6—Orientation of directional antennas at directional cell sites in AMPS.

number of cell sites needed to serve the existing telephone traffic load is the chief motivation for using directional cell sites.

In comparison with an omnidirectional land transmitting antenna, a directional antenna can deliver the same signal level in the region that it serves while causing substantially less interference within co-channel cells which lie outside the 120-degree wedge which the front lobe illuminates. Similarly, a directional land receiving antenna substantially attenuates interference received from mobile units at bearings not spanned by the front lobe. If omnidirectional systems and directional systems are to have comparable radio-frequency signal-to-interference statistics, the directional system can operate with a smaller co-channel reuse ratio, that is, a closer spacing between co-channel sites. By eq. (2), the smaller co-channel reuse ratio is equivalent to a smaller number of cells per cluster, or, more to the point, a smaller number of channel sets. Since the total number of channels is fixed, the smaller number of sets means more channels per set and per cell site. Each site can carry more traffic, thereby reducing the total number of sites needed for a given total load.

The use of three faces at each site with the orientation described above leads to certain convenient symmetries and relationships in the system design. A hexagonal cellular system could be designed, however, for a different azimuthal orientation of the directional antennas or for some other number of faces at each site.

5.2 Functional description of system operation

This section offers the reader a glimpse of the system control architecture, which is discussed in greater detail in Ref. 4. The main entities of an Advanced Mobile Phone Service system are the Mobile Telephone Switching Office (MTSO), the cell sites, and the mobile units. The central processor of the MTSO controls not only the switching equipment needed to interconnect mobile parties with the land telephone network, but also cell-site actions and even many of the actions of mobile units through commands relayed to them by the cell sites.

The MTSO is linked with each cell site by a group of voice trunks—one trunk for each radio channel installed in the site—and two or more data links, over which the MTSO and cell site exchange information necessary for processing calls. Every cell site contains one transceiver for each voice channel assigned to it and the transmitting and receiving antennas for these channels. The cell site also contains signal-level monitoring equipment and a "setup" radio, whose purpose is explained below.

The mobile equipment consists of a control unit, a transceiver, a logic unit, and two antennas. The control unit contains all the user interfaces, such as a handset, various pushbuttons, and indicator lights. The transceiver uses a frequency synthesizer to tune to any allocated

channel. The logic unit interprets customer actions and system commands and controls the transceiver and control units. A single antenna is used for transmission; two antennas together are used to provide space diversity for reception.

A few allocated radio channels serve as "setup" channels rather than voice channels; these channels are used primarily for the exchange of information needed to establish or set up calls. Applying the frequency-reuse concept to setup channels minimizes the number of channels withheld from voice use. Ordinarily, each site has one such channel. Whenever a mobile unit is turned on but the user is not engaged in a call, the mobile unit simply monitors a setup channel. The unit itself chooses which one of the various channels to monitor by sampling the signal strength on all members of a standard group of setup channels. The mobile unit then tunes to the channel which yields the strongest measurement, synchronizes with the data stream being transmitted by the system, and begins interpreting the data. Ordinarily, the mobile unit will remain on this channel; in some cases, the received data will indicate that the mobile unit should sample the signal strength on another set of channels before making a final choice. The mobile unit continues to monitor the chosen setup channel unless some condition, such as poor reception, requires that the choice of a channel be renewed. The setup-channel data words include the identification numbers of mobile units to which calls are currently being directed.

When a mobile unit detects that it is being called, it quickly samples the signal strength on all the system's setup channels so that it can respond through the cell site offering the strongest signal at the mobile unit's current position. The mobile unit seizes the newly chosen setup channel and transmits its page response. The system then transmits a voice-channel assignment addressed to the mobile unit, which, in turn, tunes to the assigned channel, where it receives a command to alert the mobile user. A similar sequence of actions takes place when the mobile user originates a call.

While a call is in progress, at intervals of a few seconds the system examines the signal being received at the serving cell site (the site that is handling the call). When necessary, the system looks for another site to serve the call. When it finds a suitable site, the system sends the mobile unit a command to retune to a channel associated with that site. While the mobile unit is changing channels, the MTSO reswitches the land party to the trunk associated with the new channel's transceiver. The periodic examination of a mobile unit's signal is known as "locating." The act of changing channels has come to be called "handoff."

The sole purpose of the locating function is to provide satisfactory transmission quality for calls.[5] In this context, the term "locating" is

really a misnomer. The term was coined in the early stages of the evolution of the cellular concept, when system designers supposed that it would be necessary to know the physical position of the mobile unit accurately.

VI. SELECTION OF KEY SYSTEM PARAMETERS

This section discusses the current recommended values for some key system geometrical parameters and the methodologies which led to them. The most important objectives in the setting of parameters are cost restraints, good transmission quality, and a large ultimate customer capacity. In some contexts, conflicts appear among these objectives, and tradeoffs must be made so that no one objective is seriously undercut to benefit another.

6.1 Cell-site position tolerance

Previous sections have alluded to perfectly regular spacing of hexagonal cell sites. In practice, however, the procurement of space for cell sites may be one of the most difficult practical hurdles in engineering and installing cellular systems.

The current design permits a cell site to be positioned up to one-quarter of the nominal cell radius away from the ideal location. The site position tolerance has far more impact on transmission quality than on cost or capacity. Consequently, an analysis was made of the effect of the cell-site position tolerance on the overall probability distribution of the RF signal-to-interference ratio (S/I) on voice channels in mature systems. For simplicity and concreteness, the analysis focused on the value of RF S/I ratio which falls at the tenth percentile of the overall S/I distribution. This level decreased gradually as the cell-site position tolerance increased from 0 to about one-fourth of a cell radius, but it decreased rapidly beyond this break point. The tolerance was therefore set at a quarter radius to allow system administrators as much leeway as possible in positioning sites without significantly degrading the transmission quality.

6.2 Maximum cell radius

Setting the maximum cell radius, which is to be used in a system at its inception, is part of the general problem of achieving a satisfactory compromise between the objectives of low cost and good transmission quality. The maximum cell radius has only an indirect effect on the system objective of a large ultimate capacity.

Transmitter power is another important element in the overall reconciliation of low cost with high-quality transmission. In present-day mobile-telephone systems, the transmitter power of mobile units is smaller by an order of magnitude than that of the land stations. To

provide adequate reception of mobile transmissions emanating from any place where mobile units receive land transmissions, satellite receiver sites are deployed throughout the coverage area. A cellular system might be designed in this manner. However, the AMPS system designers consider a "balanced" system with comparable transmitter power in mobile units and cell sites to be a more economical design, because satellite receiver sites would represent a substantial fraction of the cost of full-fledged cell sites, yet possess far fewer capabilities.

When a system is first established, there is normally little frequency reuse. Since each initial cell is relatively large, the total number of cells needed to span the desired coverage area does not greatly exceed the number of channel sets into which the total allocation is partitioned. Even though two or more cell sites may be assigned the same channel set, mutually exclusive subsets can be used in the co-channel sites until a certain amount of growth in telephone traffic has occurred. For an initial period, therefore, the main channel impairment to contend with is ambient noise, both inevitable receiver thermal noise and man-made environmental noise.

In a startup system, an increase in land and mobile transmitter power, all other system parameters being held constant, would improve transmission quality by raising RF signal-to-noise (S/N) ratios, but it would also raise the system cost. From a broader perspective, however, increased transmitter power could be used to reduce system cost rather than to improve transmission quality. Increased power would permit the use of a larger initial cell radius for the same level of transmission quality, which in turn would allow fewer cell sites to cover the desired area. If an extra expenditure on transmitter power yields a greater cost saving in cell-site construction, the expenditure is desirable. At some level, however, additional transmitter power ceases to pay for itself. Not only does the incremental cost per decibel mount, but eventually practical considerations of feasibility and reliability also enter in. Furthermore, the relatively high level of transmitter power that is beneficial in the startup phases of cellular systems is largely superfluous in mature phases, since each stage of cell splitting essentially halves the mean distance between mobile units and their serving cell sites.

The chosen value of 10 watts delivered to the transmitting antennas is based on an evaluation of the cost, reliability, and power drain of present-day transmitters in the 800- to 900-MHz range. To supply approximately 10 watts of power at the antenna terminals, the system design requires 12 watts from mobile transmitters and 40 watts from cell-site transmitters to compensate for cable and combiner losses.

Cell-site antenna elevation and gain (in any vertical plane) influence the tradeoff between cost and transmission quality in much the same way as transmitter power. As with transmitter power, the selected

figures for antenna gain and elevation are the largest values normally achievable without excessive costs. The expected range of antenna gain is 6 to 8 dB relative to a dipole; the expected range of elevation above the ground, 100 to 200 feet.

Assuming that transmitter power and cell-site antenna gain and elevation are already established, the tradeoff between cost and quality in the early stages of AMPS growth is governed by the value chosen for the cell radius. Since increasing the radius both decreases cost and degrades transmission quality, the transmission-quality objective allows a controlled level of imperfection for the sake of economy.

The sound quality of AMPS calls is intended to be comparable in acceptability to the sound quality on calls over the land telephone network, but setting system parameters requires that this general guideline be reduced to more concrete terms.

The maximum cell radius depends on both subjective and statistical factors. To meet the sound-quality objective, designers required information both on customer opinions of mobile-telephone channels at 800 to 900 MHz and on the propagation of energy at these frequencies. In a subjective testing program,[6] subjects rated the quality of simulated and actual mobile-telephone channels subjected to the rapid Rayleigh fading encountered in UHF mobile communications. The test results showed that at an RF S/N ratio of 18 dB, most listeners considered the channel to be good or excellent. The system designers concluded that the S/N ratio in a working system should exceed 18 dB with high probability. The AMPS transmission-quality objective was therefore quantified for design purposes as a requirement that this S/N value be exceeded in 90 percent of the area covered by any system.

Our knowledge of propagation is based largely on an extensive measurement program performed by Bell Laboratories in Philadelphia in the early 1970s and in Newark more recently.[7] These measurement results corroborate studies performed in Tokyo,[8] New York,[9] and suburban areas of New Jersey.[10] All these investigations and others[11] show that, for a given distance r between transmitter and receiver, the probability distribution of the path loss (attenuation) in decibels is approximately Gaussian. The mean of the distribution (in decibels) is approximated by a function of the form $k + 10n \log_{10} r$, in which k is a constant for a given transmitter-receiver pair and n is known as the path-loss exponent. The standard deviation amounts to several decibels. The bodies of data associated with different transmitter sites yield different numerical results for all these parameters, but the values which emerge from the total ensemble of Philadelphia and Newark measurements are a path-loss exponent n on the order of 4 and a standard deviation of roughly 8 dB.

At Bell Laboratories, a computer simulation was written to predict

many aspects of the behavior of a cellular system, including the overall statistics of received signal level. This simulation incorporates a propagation model based on the Philadelphia measurements, and it also models vehicle movements and the details of the locating algorithm.

The results of this simulation show that in an environment similar to Philadelphia a cell radius of 8 miles allows the system to meet the requirement that the S/N ratio be above 18 dB over 90 percent of the coverage area. Somewhat different values of maximum cell radius may be appropriate for situations in which any of the relevant parameters, such as the path-loss exponent, environmental noise, antenna gain, or antenna elevation, differ substantially from the assumed values.

6.3 Minimum cell radius

In the AMPS system, additional cell sites needed to relieve the telephone traffic demand on existing sites will be positioned midway between adjacent old sites. This simple procedure cuts the distance between adjacent sites in half and therefore cuts the cell radius by a factor of 2 and the cell area by a factor of 4.

The minimum cell radius, which is the cell radius after the final stage of cell splitting, has little effect on the system cost per customer or on transmission quality, but it plays a vital part in setting the ultimate system capacity. Each stage of cell splitting multiplies the number of cell sites in the desired coverage area by a factor of about 4. The system's total traffic-carrying capacity is also increased by essentially the same factor. In principle, the cell-splitting process could be repeated an indefinite number of times, but system designers cite a 1-mile cell radius as a practical minimum. If start-up cells have an 8-mile radius, three stages of cell splitting are possible. There is no insurmountable physical barrier to having smaller cells, but the greatest practical obstacles are the cell-site position tolerance and the burden of frequent handoffs. As previously stated, cell sites should be positioned within a quarter of a cell radius of their ideal locations. A tolerance of a quarter mile (corresponding to a 1-mile radius) is probably the most stringent requirement that can be contemplated. The mean distance traveled between handoffs is bound to decrease as the cell radius decreases. In a system composed of many cells with a radius of much less than one mile, handoffs would consume a significant fraction of the MTSO central processor's capacity.

6.4 Co-channel reuse ratio

The discussion of directional cell sites explained the economic incentive for minimizing the ratio of D, the distance between co-channel cell sites, to R, the cell radius. The co-channel reuse ratio (D/R) also has an impact on both the transmission quality and the ultimate customer capacity of the system. The influence on transmission quality

arises because the D/R ratio materially affects co-channel interference statistics. Since this ratio determines the number of channels per channel set, it sets a limit on each site's traffic-carrying capacity, which in turn limits the ultimate system capacity.

The allowable minimum value of D/R was set in much the same way as the maximum cell radius. Making D/R as small as possible serves the objectives of low cost and large capacity. On the other hand, making D/R as large as possible benefits transmission quality. As in the determination of the maximum cell radius, a compromise among objectives is achieved through the kind of transmission-quality objective described in the discussion of maximum radius.

The subjective testing program mentioned previously included an evaluation of the effect of co-channel interference on listeners' opinions. The results indicated that most of the subjects considered the transmission quality of a channel to be good or excellent at an S/I of 17 dB. To satisfy the AMPS quality objective, a system must provide an S/I of 17 dB or greater over 90 percent of its coverage area. The system simulation mentioned above shows that, in an environment similar to Philadelphia or Newark, a system meets the S/I requirement if the separation between co-channel sites is 4.6 cell radii when 120-degree directional antennas are used and 6.0 cell radii when omnidirectional antennas are used. These co-channel reuse ratios correspond, by eq. (2), to 7 cells per cluster (equivalently 7 disjoint channel sets) for directional sites and 12 cells per cluster for omnidirectional sites.

VII. DEFINITION AND DEPLOYMENT OF CHANNEL SETS

A complete description of a plan for deploying channels in a coverage area requires that some additional facts and procedures be specified. The degree of foresight with which channel sets are defined and used can materially affect the system's transmission quality, cost, and ease of adaptation to growth in telephone traffic.

7.1 Reduction of adjacent-channel interference

The design of a mobile-telephone system must include measures to limit not only co-channel interference but also adjacent-channel interference. Although the IF filters of both the cell-site and mobile-unit receivers significantly attenuate signals from the channels adjacent in frequency to the desired channel, it is advisable to avoid circumstances in which the received level of an adjacent channel greatly exceeds that of the desired channel.

This situation would arise at a cell site, for example, if one mobile unit were many times farther away from its serving cell site than another mobile unit being served by the same site on an adjacent channel. With a distance ratio of 10, for instance, the received level of the adjacent channel at the cell site could easily be 40 dB higher than

the level of the desired channel. In the presence of fading, severe adjacent-channel interference would result unless the receiver IF filter could greatly attenuate the adjacent channel. In general, a substantial spectral guard band would be required between channels to permit IF filters to reject the interference adequately.

Fortunately, in a cellular system, since only a fraction of the allocated channels belong to any one channel set, it is possible to avoid the use of adjacent channels in the same cell site, thereby keeping the probability of severe adjacent-channel interference low. Since stringent IF attenuation of adjacent channels is not essential, no guard band is needed.

AMPS voice channels have an FM peak deviation of 12 kHz and a spacing of 30 kHz.[12] With this spacing, 666 duplex channels can be created out of a 40-MHz spectrum allocation. The use of adjacent channels at the same site would require a larger channel spacing, and fewer channels would be available from the allocation. In the AMPS system, the largest possible frequency separation is maintained between adjacent members of the same channel set. Suppose that channels are numbered sequentially from 1 upward and that the frequency difference between channels is proportional to the algebraic difference of their channel numbers. If N disjoint channel sets are required, the nth set ($1 \leq n \leq N$) would contain channels n, $n + N$, $n + 2N$, etc. For example, if $N = 7$, set 4 would contain channels 4, 11, 18, etc.

In some cases, system designers can also prevent a secondary source of adjacent-channel interference by avoiding the use of adjacent channels in geographically adjacent cell sites. Figure 7 shows a cell-site pattern for 12 disjoint channel sets (12 cells per cluster). This pattern can be used in startup phases of AMPS, during which all sites will be equipped with omnidirectional voice-channel antennas. Only sets with adjacent set numbers (including 12 and 1) contain any adjacent channels. In the figure, each site is labeled with the number of its channel set. Center-excited cells are drawn in to aid in visualizing the nominal area in which each set is most likely to be used.

As previously discussed, when 120-degree directional cell-site antennas are employed in AMPS, transmission-quality considerations call for seven cells per cluster. In this case, it is impossible to avoid having adjacent channels at adjacent sites, because if there are only seven channel sets, any site plus its six neighbors constitute a complete cluster in which every channel may be assigned exactly once. With 120-degree directional antennas, however, it is possible to subdivide the seven channel sets and deploy the subsets geographically in such a way that the received adjacent-channel interference, at both the mobile units and cell sites, is usually attenuated by the front-to-back ratio of the cell-site directional antennas. (An example below illustrates this effect.) The AMPS plan subdivides each of the seven channel sets

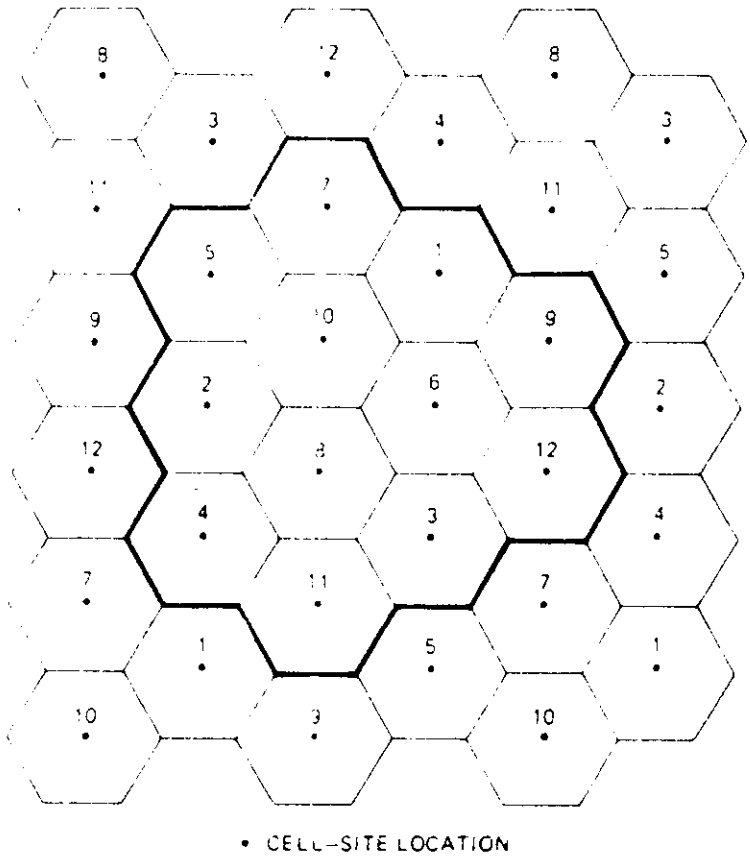


Fig. 7—Channel-set deployment pattern for 12 cells per cluster.

into three subsets. For example, set 4, containing channels 4, 11, 18, 25, 32, 39, 46, 53, 60, etc., subdivides into subset 4a with channels 4, 25, 46; subset 4b with channels 11, 32, 53, etc.; and set 4c with channels 18, 39, 60, etc. The notation is simplified if we simply number the channel subsets from 1 to 21, so that set n is subdivided into subsets n, $n + 7$, and $n + 14$.

Figure 8 shows one acceptable pattern for assigning channel subsets to cell-site faces. (The appendix to this paper describes a simple algorithm for assigning channel sets to cells; this algorithm would produce a differently labeled but equally acceptable pattern.) In Fig. 8, corner-excited cells are shown whose sides are projections of the edges of the antennas' 120-degree front lobes. Subsets with sequential subset numbers contain adjacent channels and are assigned to faces in such a way that they do not cover the same corner-excited cell, even though they may reside in adjacent sites. This procedure attenuates adjacent-channel interference by the cell-site antenna's front-to-back ratio in situations which otherwise would cause problems. For instance, in Fig. 8, suppose that a channel of subset 6 is serving the mobile unit at point *M*. The adjacent-channel interference that exists in both directions between the mobile unit and the cell site using subset 7 is attenuated by the front-to-back ratio of the directional antennas which transmit and receive subset 7.

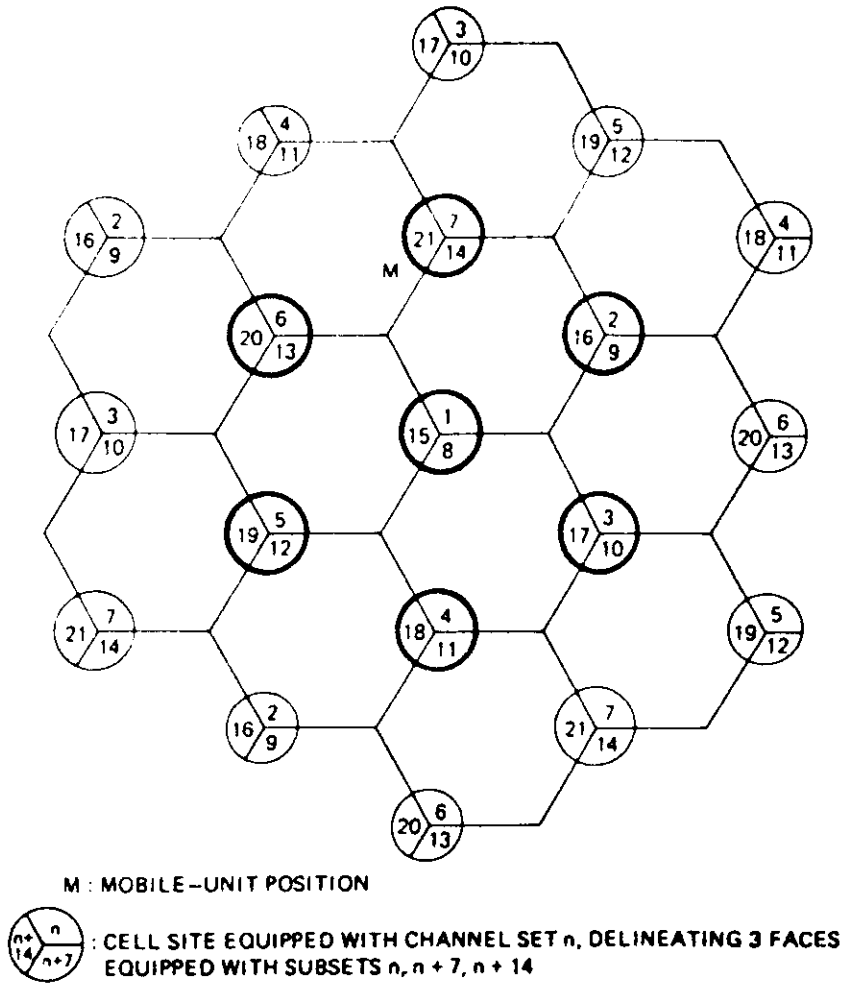


Fig. 8—Channel-subset deployment pattern for seven cells per cluster and three faces per cell site.

7.2 Cell splitting

The practical significance of cell splitting is that the distance between adjacent cell sites is cut in half, and through the action of the locating algorithm, the nominal coverage area of newly established cell sites is reduced to a quarter of the nominal area previously covered by existing sites. Conversely, wherever cell splitting occurs, it quadruples the cell-site density. The AMPS cell-splitting plan sets the ideal location for new sites at points midway between neighboring existing sites, although the actual position may be anywhere within a distance of one-quarter of a (smaller) cell radius. The previously existing cell sites together with the new ones form a hexagonal cellular lattice.

In the transition from a system based on 12 channel sets and omnidirectional cell-site antennas to one based on 21 channel subsets and directional antennas, a gradual alteration may be necessary in the assignment of channel frequencies to cell sites. Once directional operation is established, however, splitting does not cause any further alteration of existing channel assignments. Figure 9 shows an array of directional cell sites identified by single set numbers. In one area, six new cell sites have been established. The channel set assigned to any

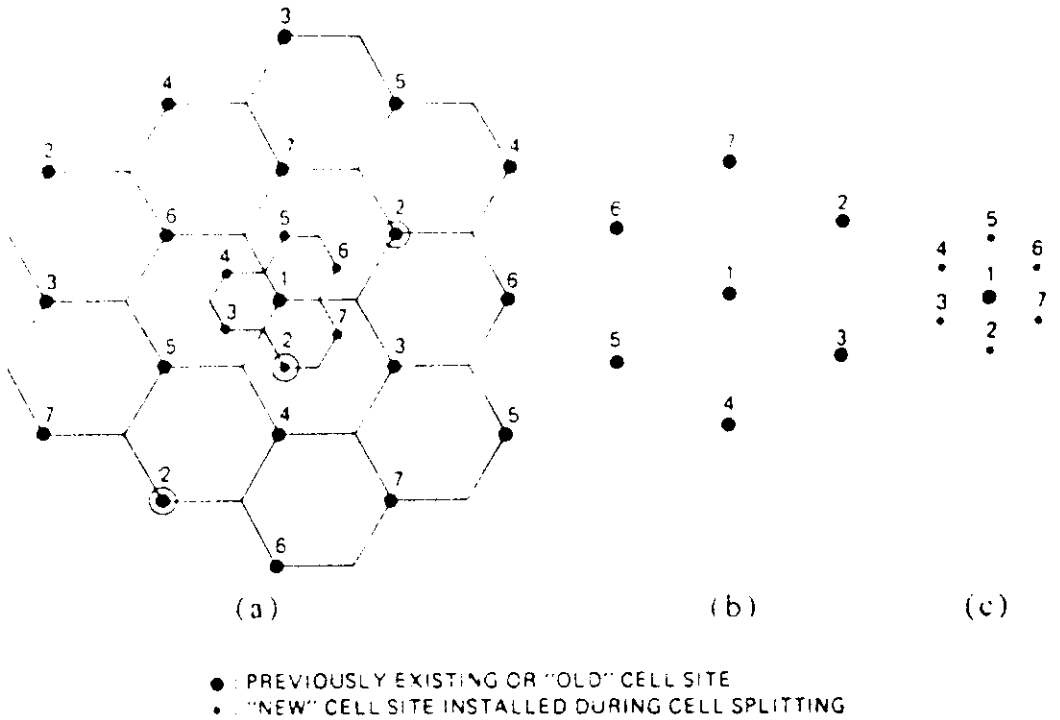


Fig. 9—Channel-set deployment pattern for seven cells per cluster, with multiple cell sizes. (a) Overall cell-site pattern. (b) Orientation of cluster of cell sites in larger-cell pattern. (c) Orientation of cluster of cell sites in smaller-cell pattern.

new site is determined by observing that the new site lies midway between two co-channel sites, each one situated a little more than one (larger) cell diameter away from the new site, both of which use the same channel set. This is the channel set which should be assigned to the new site. The sites labeled 2 in Fig. 9a are circled to illustrate this geometrical relationship. Figures 9b and 9c extract from Fig. 9a the channel-set patterns of the cell-site clusters at different stages of cell splitting. Successive stages of cell splitting preserve the internal geometrical relations within the cluster, but each stage causes the cluster to be rotated counterclockwise 120 degrees.

7.3 The overlaid-cell concept

In a coverage area where two or more sizes of cells exist simultaneously, special care must be taken to guarantee the correct minimum distance D between cell sites equipped with the same voice channels. As previously discussed, a certain co-channel reuse ratio D/R must be maintained, but when a system includes multiple cell sizes, the radius R has different values for different sites.

Figure 10 illustrates some unusual situations that arise whenever groups of cells of different sizes abut. Site A_1 lies within a group of larger cells. For a cellular pattern of seven cells per cluster, the nearest co-channel sites within the group of larger cells, such as sites A_2 and A_3, should be separated from site A_1 by a distance of 4.6 larger-cell radii. Within the group of smaller cells, co-channel sites such as sites A_4 and A_5 are separated from each other by a distance of 4.6 smaller-

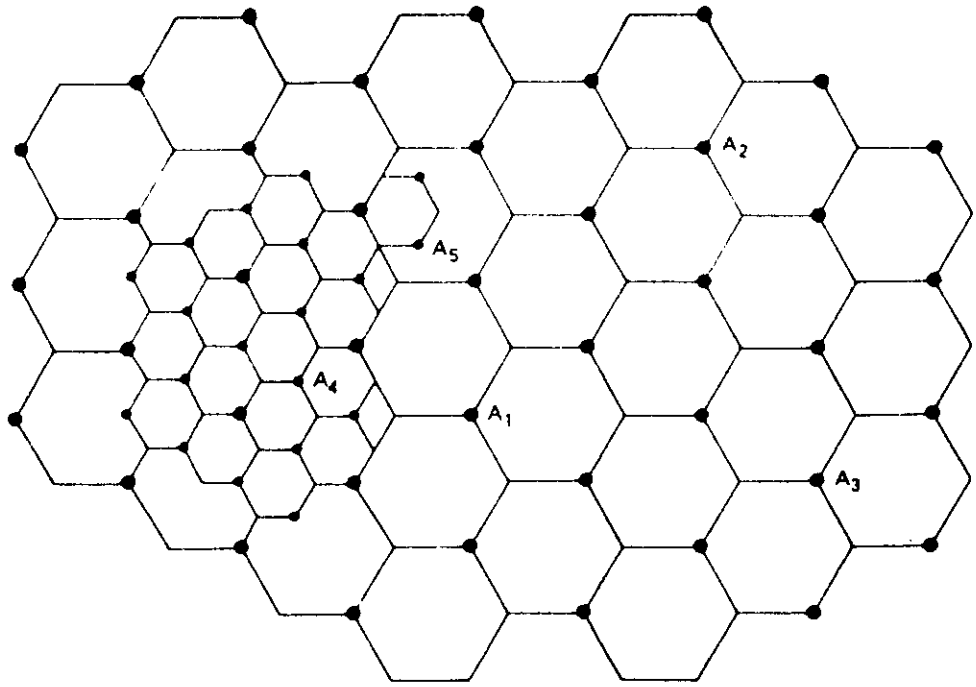


Fig. 10—Illustration of mixed cell sizes for discussion of overlaid-cell concept.

cell radii, or 2.3 larger-cell radii. Site A_1 is also in the correct position to be a nearest co-channel neighbor of sites A_4 and A_5. The D/R ratio is satisfied for sites A_4 and A_5 because the appropriate value of R for these sites is the smaller-cell radius. Channels installed in site A_1 would cause no undue co-channel interference on calls served by sites A_4 and A_5, because a mobile unit being served by one of these latter sites would tend to be within a range of about one smaller cell radius.

The troublesome questions pertain to calls served by site A_1. If this site is to serve a larger-cell area, then it cannot use any of the same channels as sites A_4 and A_5, because the co-channel reuse ratio (D/R) would not be satisfied for site A_1 if the larger-cell radius is taken as the appropriate value for R. This ratio would be satisfied if the smaller-cell radius could somehow be made applicable to site A_1, but restricting site A_1 to serving only a smaller-cell area could mean inadequate coverage for some areas further removed from the site.

The dilemma affecting site A_1 is resolved by invoking the overlaid-cell concept. This concept recognizes that, when multiple cell sizes co-exist, the cellular pattern is best viewed as the superposition of a fragmentary smaller-cell pattern on top of a complete larger cell pattern. The underlying larger-cell pattern does not disappear in a given region until the overlaid smaller-cell pattern is complete in that region.

Implementing the overlaid-cell concept requires that, in a region where cells of two sizes are present, the channel subset assigned to any cell-site face must be further subdivided into a larger-cell group and a

smaller-cell group. Each face of an older, previously existing site will use some of its channels to continue coverage of the same larger cellular area as before. The remainder of the channels assigned to the face will be restricted to covering a smaller area, corresponding to the smaller cell size. The subdivision of a subset into larger- and smaller-cell groups for an existing site is governed by the channel requirements of its new co-channel neighbors. For example, in Fig. 10, any channel installed in site A_4 or A_5 must be restricted to smaller-cell use in site A_1. The way that a channel is restricted to smaller-cell use is simply to reassign the channel in software to a channel group which is treated as if it were serving a smaller cell. When appropriate, a call being served by a smaller-cell channel will be handed off to a neighboring new site if there is one, or otherwise to a channel belonging to the larger-cell group of the same face on which the call is already being served. As the telephone traffic loads grow in the new sites, reassignment of more and more channels in the old site to a smaller-cell group must follow, thereby reducing the capacity of the older site to serve the larger-cell area. For this reason, not only the local growth in telephone traffic around any older site but also the growth in traffic carried by that site's new co-channel neighbors can force cell splitting around the older site.

The various procedures described in this section for channel-set definition and deployment and for cell splitting allow the system to grow gradually and, on the whole, gracefully in response to a growing telephone traffic load. When an existing site reaches its traffic-carrying capacity, new sites are added around it one by one, only as needed, while the older site makes a gradual transition from larger-cell operation to smaller-cell operation.

VIII. SUMMARY

The FCC's allocation of a relatively large block of spectrum for public mobile communications has made a large-scale, economical mobile-telephone service feasible. The need for a method of serving many thousands of customers in a single local coverage area, while using a limited spectral allocation equivalent to several hundred voice channels, has spurred the evolution of the cellular concept. In practical terms, a cell is the area in which one particular group of channels is more likely to be used for mobile telephone calls than any other group. The essential elements of the cellular concept, frequency reuse and cell splitting, allow a cellular system to use spectrum efficiently, to grow gradually, and to supply service in response to a geographical pattern of demand.

In the AMPS system, the lattice of cell sites is designed to create a pattern of hexagonal cells. In the initial growth phase of a system in a given locality, cell-site voice-channel transmitters and receivers con-

nect to omnidirectional antennas. In later stages, cell sites have three faces, equipped with 120-degree directional antennas. The omnidirectional plan minimizes the startup costs for new systems, whereas the directional plan confines costs in mature systems by reducing the total number of sites required to serve a given offered load.

The key geometrical parameters of AMPS were chosen primarily to satisfy the objectives of moderate cost, good transmission quality, and a large ultimate customer capacity. In some cases, tradeoffs must be made among these joint objectives.

The ways in which channel sets are defined and distributed among cell sites in AMPS keep co-channel and adjacent-channel interference within acceptable bounds. The procedure for assigning channels to new cell sites introduced during cell splitting promotes graceful growth in response to increasing demand. As new sites are added, previously existing sites make a gradual transition from larger-cell operation to smaller-cell operation.

IX. ACKNOWLEDGMENTS

During three decades, many people have contributed to the definition of a mobile-telephone system structure capable of realizing the promises of the cellular concept. We shall mention just a few. W. R. Young and J. S. Engel helped to crystallize the overall system objectives. The cellular geometry of AMPS and the techniques for deploying and utilizing channel frequencies largely reflect the contributions of R. H. Frenkiel and P. T. Porter. In particular, the overlaid-cell concept, which governs system growth procedures, is the work of R. H. Frenkiel. M. A. Castellano also worked out many of the geometrical details of channel-set deployment. J. A. O'Brien and G. D. Ott contributed to the iterative process of proposing and evaluating the details of the locating algorithm.

APPENDIX

Fundamentals of Hexagonal Cellular Geometry

Certain intriguing mathematical relations emerge when one deals with hexagonal cellular geometry, yet there appears to be no published summary of all the basic relations with explanations of how they arise. This appendix is intended to fill the gap. We also present a novel algebraic method for using the coordinates of a cell's center to determine which channel set should serve the cell.

Figure 11 shows the most convenient set of coordinates for hexagonal geometry. The positive halves of the two axes intersect at a 60-degree angle, and the unit distance along either axis equals $\sqrt{3}$ times the cell radius, the radius being defined as the distance from the center of a cell to any of its vertices. With these coordinates, an array of cells can

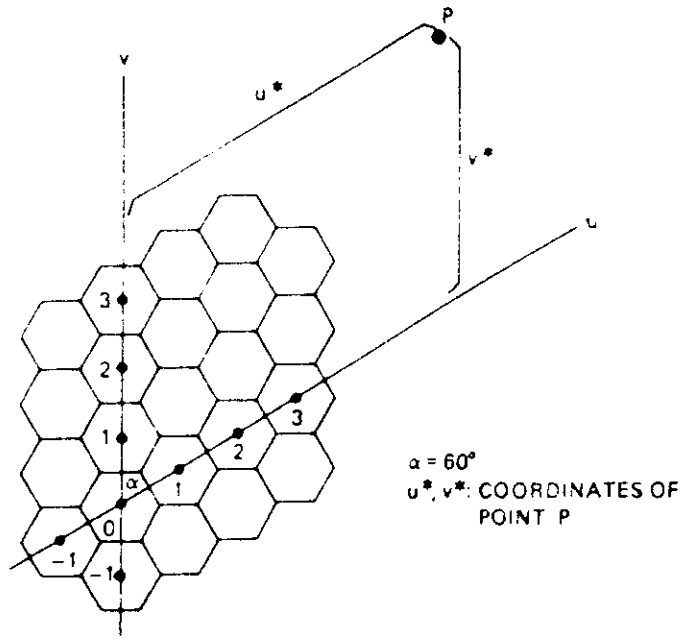


Fig. 11—A convenient set of coordinates for hexagonal cellular geometry.

be laid out so that the center of every cell falls on a point specified by a pair of integer coordinates.

The first useful fact to note is that, in this coordinate system, the distance d_{12} between two points with coordinates (u_1, v_1) and (u_2, v_2), respectively, is

$$d_{12} = \sqrt{(u_2 - u_1)^2 + (u_2 - u_1)(v_2 - v_1) + (v_2 - v_1)^2}. \tag{3}$$

Using this formula we can verify that the distance between the centers of adjacent cells is unity and that the length of a cell radius R is

$$R = 1/\sqrt{3}. \tag{4}$$

We can calculate the number of cells per cluster, N, by some heuristic reasoning. The directions given in Section IV of this paper for locating co-channel cells result in a co-channel relationship between the reference cell with its center at the origin and the cell whose center lies at $(u, v) = (i, j)$, where i and j are the integer "shift parameters," with $i \geq j$. (See Fig. 4 for an illustration.) By eq. (3), the distance D between the centers of these or any other nearest neighboring co-channel cells is

$$D = \sqrt{i^2 + ij + j^2}. \tag{5}$$

Figure 4 illustrates the universal fact that any cell has exactly six equidistant nearest neighboring co-channel cells. Moreover, the vectors from the center of a cell to the centers of these co-channel cells are separated in angle from one another by multiples of 60 degrees. These same observations also hold for any arbitrary cell and the six cells immediately adjacent to it. The idea presents itself to visualize each

cluster as a large hexagon. In reality, the cluster, being composed of a group of contiguous hexagonal cells, cannot also be exactly hexagonal in shape, but it is nevertheless true that a properly visualized large hexagon can have the same area as a cluster. For proper visualization, refer to Fig. 12. The seven cells labeled A are reproduced from Fig. 4. The center of each A cell is also the center of a large hexagon representing a cluster of cells. Each A cell is imbedded in exactly one large hexagon, just as it is contained in exactly one cluster. All large hexagons have the same area, just as all clusters have the same area. The large hexagons cover the plane with no gaps and no overlaps, just as the clusters do. We therefore claim that the area of the large hexagon equals the area of any valid cluster. This area can be deduced from results already presented. We noted above that the distance between the centers of adjacent cells is unity. By eq. (5), the distance between centers of the large hexagons is $\sqrt{i^2 + ij + j^2}$.

Consequently, since the pattern of large hexagons is simply an enlarged replica of the original cellular pattern with a linear scale factor of $\sqrt{i^2 + ij + j^2}$, then N, the total number of cell areas contained

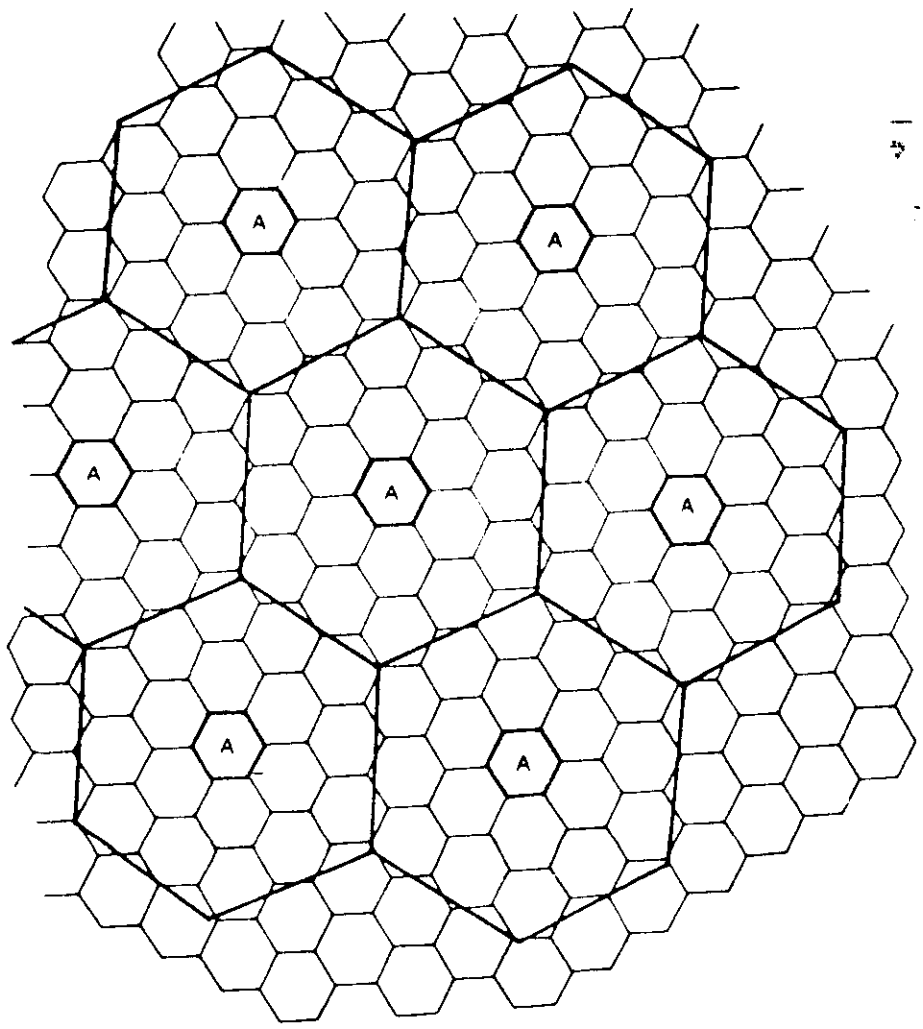


Fig. 12—Illustration for the heuristic determination of the number of cells per cluster.

in the area of the large hexagon, is the square of this factor, namely

$$N = i^2 + ij + j^2. \tag{6}$$

By combining eqs. (4), (5), and (6), we obtain the classical relationship between the co-channel reuse ratio D/R and the number of cells per cluster N:

$$D/R = \sqrt{3N}. \tag{7}$$

The rather cumbersome procedure described in the main body of this paper for performing a cellular layout can be replaced by a simple algebraic algorithm in certain cases of practical interest, namely those cases in which the smaller shift parameter j equals unity. (A pattern of 7 cells per cluster falls into this category.) For these cases, it is convenient to label the cells by the integers 0 through N-1. Then the correct label L for the cell whose center lies at (u, v) is given by

$$L = [(i + 1)\, u + v] \bmod N. \tag{8}$$

The application of this simple formula causes all cells which should use the same channel set to have the same numerical label.

REFERENCES

1. W. R. Young, "AMPS: Introduction, Background, and Objectives," B.S.T.J., this issue, pp. 1–14.
2. Richard H. Frenkiel, "A High-Capacity Mobile Radiotelephone System Model Using a Coordinated Small-Zone Approach," IEEE Trans. Veh. Technol., *VT-19*, No. 2 (May 1970), pp. 173–177.
3. Bell Laboratories, "High-Capacity Mobile Telephone System Technical Report" (December 1971), submitted to FCC.
4. Z. C. Fluhr and P. T. Porter, "AMPS: Control Architecture," B.S.T.J., this issue, pp. 43–69.
5. Gary D. Ott, "Vehicle Location in Cellular Mobile Radio Systems" IEEE Trans Veh. Technol., *VT-26*, No. 1 (February 1977), pp. 43–46.
6. V. H. MacDonald, "Subjective Voice-Circuit Quality Measurements in the Presence of Multipath Fading," Symposium on Microwave Mobile Communications, Boulder, Colo., 1972.
7. G. C. DiPiazza, A. Plitkins, and G. I. Zysman, "AMPS: The Cellular Test Bed," B.S.T.J., this issue, pp. 215–248.
8. Y. Okumura, E. Ohmori, T. Kawano, and K. Fukuda, "Field Strength and its Variability in VHF and UHF Land Mobile Service," Rev. Elec. Comm. Lab., *16* (September 1968), pp. 825–873.
9. W. Rae Young, Jr., "Comparison of Mobile Radio Transmission at 150, 450, 900, and 3700 MC," B.S.T.J., *31*, No. 6 (November 1952), pp. 1068–1085.
10. D. C. Cox, "Time-and-Frequency Domain Characterizations of Multipath Propagation at 910 MHz in a Suburban Mobile-Radio Environment," Radio Sci., *7* (December 1972), pp. 1069–1077.
11. W. C. Jakes, Jr., ed., *Microwave Mobile Communications*, New York, John Wiley, 1974.
12. G. A. Arredondo, J. C. Feggler, and J. I. Smith, "AMPS: Voice and Data Transmission," B.S.T.J., this issue, pp. 97–122.

Land Mobile Radio Systems—a Tutorial Exposition

S.C. Gupta
R. Viswanathan
R. Muammar

An in-depth tutorial on land mobile radio systems

June 1985—Vol. 23, No. 6
IEEE Communications Magazine

IN MOBILE RADIO SYSTEMS, the propagation between the transmitting antenna and the mobile unit antenna is over several paths, namely, the line of sight path and the paths due to scattering caused by reflections from and diffractions around obstructions. These interfering signals produce a complex standing-wave pattern of varying field strength, with maxima and minima being of the order of a quarter wavelength apart. As a result of the vehicle movement through this standing wave pattern, the received signal experiences random variations in both amplitude and phase. Fades of 40dB or more below the mean signal level are common, with successive minima occurring about every half wavelength of the carrier transmission frequency. The received signal fluctuates as the vehicle moves, thus distorting speech when transmitted by conventional methods.

Starting from a model based on multipath wave interference arising from multiple scattering of the waves by buildings and other structures in the vicinity of the mobile unit, Clarke [1] and Gans [2] have shown that the envelope of the mobile radio signal is Rayleigh distributed when measured over distances of a few tens of wavelengths, where the mean signal is sensibly constant, whereas, the phase of the received signal is uniformly distributed from 0 to 2π. Therefore, the probability density function of the received signal envelope S relative to the local mean $\bar{a} = E(S)$, can be written as:

$$P(S/\bar{a}) = \frac{\pi S}{2\,\bar{a}^2} \exp\left(-\frac{\pi S^2}{4\,\bar{a}^2}\right), \quad S > 0$$

Although the statistics of the received signal envelope are Rayleigh distributed, the local-mean $\bar{a}$ varies typically between 6 to 12 dB due to shadowing. Shadowing of the radio signal by buildings and hills leads to a gradual change in the local-mean which can be characterized statistically by log-normal distribution [3] with two parameters (m_d, σ). The log-normal distribution with parameters (m_d, σ) in dB is described by the probability density function:

$$P(\bar{a}) = \frac{K}{\bar{a}\,\sigma} \exp\left(-\frac{1}{2\sigma^2}\{20 \log_{10} \bar{a} - m_d\}^2\right)$$

where K is a constant, m_d and σ^2 are the mean and variance of the corresponding normal distribution.

The probability density function of the signal envelope under Rayleigh fading and log-normal shadowing has been derived in [4] and it is given by:

$$P(S) = \frac{k}{\sigma} \int_{-\infty}^{\infty} \frac{S}{10^{a/10}} \exp\left[-\frac{\pi S^2}{4\times 10^{a/10}}\right] \exp\left[-\frac{(a-m_d)^2}{2\sigma^2}\right] da$$

where k is a constant, and $a = 20 \log_{10} \bar{a}$

The expected rate at which the envelope crosses a specified value R, can be obtained using [5]:

$$N_R = \int_0^{\infty} \dot{S}\, P(R,\dot{S})\, d\dot{S}$$

where N_R is the level crossing rate, $\dot{S}$ is the time derivative of the envelope amplitude S, and $P(R,S)$ is the joint density function of the signal envelope at $S = R$ and $\dot{S}$. It has been shown (for Rayleigh fading only) [3] that the deeper the fade, the less frequently it is expected to occur. The average duration of fade is defined as:

$$\bar{\tau} = \frac{P(S \le R)}{N_R}$$

Both the rate N_R and the average duration $\bar{\tau}$ of the fades are useful in the evaluation of the receiver performance.

0163-6804/85/0600-0034 $01.00 © 1985 IEEE

Besides the random variations in signal amplitude and its phase, each component plane wave has a Doppler shift associated with it depending on the mobile speed, the carrier frequency, and the angle its propagation vector makes with the mobile velocity vector.

$$f_D = \frac{v}{\lambda} \cos \alpha$$

where $\frac{v}{\lambda}$ is the maximum Doppler shift ($\alpha = 0$) at the carrier wavelength λ, and v is the vehicle speed.

The signal power falls off (attenuates) rapidly as the vehicle moves away from the base station. Attenuation of the signal power with distance changes from an inverse cubic law to an inverse fourth power law. The received signal power is a function of base station and mobile unit antennas height, separation between the two antennas, transmission frequency, width and orientation of the streets in urban areas, and atmospheric conditions. Described below are some of the parameters that are useful in the evaluation of channel characteristics.

Coherence Distance

This is the minimum distance between two points for which the signals are not strongly correlated, that is with a correlation coefficient of less than 0.5 [3]. The coherence distance is typically of the order of one-half wavelength in an urban location. This property of the channel enables space diversity to be used in order to combat the effects of Rayleigh fading.

Coherence Bandwidth

The different path lengths for the received signal give rise to different propagation time delay, typical spreads in time delays range from a fraction of a microsecond to many microseconds, depending on the type of environment. The existence of the different time delays in the various waves that make up the total field, causes the statistical properties of two signals of different frequencies to become essentially independent if the frequency separation is large enough. The coherence bandwidth B_c is defined [6] as the frequency spacing between two signals with a correlation coefficient of 0.5 or less. B_c is typically from 30 kHz to 1 mHz. The frequency selective properties of mobile radio channel make it possible for some system plans to employ frequency diversity in order to combat the effects of fading.

Coherence Time

The difference in time between two samples of the signal with a correlation coefficient of 0.5 is termed the coherence time of the channel. A typical value of T_c is 1.3 ms or more. The frequency dispersive properties of the channel can be utilized in a time diversity scheme to combat the effect of fading.

The next section discusses the concepts of cellular systems and how they could provide effective service to mobile users.

Cellular Systems

A spectrum efficient high-capacity system with a flexibility to accommodate the increased user densities, called the "Advanced Mobile Phone Service" system, has been studied by MacDonald [7]. The service trial of the Advanced Mobile Phone Service (AMPS) was begun in the Chicago area in 1979. The system objectives are [8]:

1) Large subscriber capacity
2) Efficient use of the spectrum
3) Nationwide compatibility
4) Widespread availability
5) Adaptability to traffic density
6) Service to vehicles and portables
7) Regular telephone service and special services, including dispatch
8) High quality of service and affordability

Frequency reuse and cell splitting summarize the essential features of the cellular concept.

Basics of the Cellular Concept

The total coverage area is divided into interlocking polygons called a cell. Each one contains its own land radio equipment for transmission to and reception from mobile units within the cell. A cellular system could be designed with square or equilateral triangular cells, but for economic reasons, the regular hexagonal shape has been adopted for Advanced Mobile Phone Service (Fig. 1). Each cell is served by a base station located at the center of the cell or at the alternate corners of the hexagons. In the first case, the cell site base station employs omnidirectional antennas to communicate with the surrounding mobiles, while in the second case the cell site uses directional antennas with 120° beamwidth to illuminate portions of the three adjacent cells which meet at the cell site. The center-located cell concept is likely to be applied in small cities because it has the economic advantage of requiring fewer cell sites. However, in large high rise cities, the corner excitation arrangement is more practical because it gives a form of space diversity that could improve the system performance in the presence of lognormal shadowing.

A fixed number of radio channels are allocated to each cell. Since each base station provides coverage only over one cell, the group of channels allocated to a cell can be used by another cell when the two are suitably separated geographically. This is called the frequency reuse which is the second essential feature of the cellular concept. The idea of employing frequency reuse in mobile-telephone service on a shrunken geographical scale hints at the cellular concept. Instead of covering an entire local area from one land transmitter site with high power at a high elevation, the service provider can distribute transmitters of moderate power throughout the coverage area, thus, increasing the system

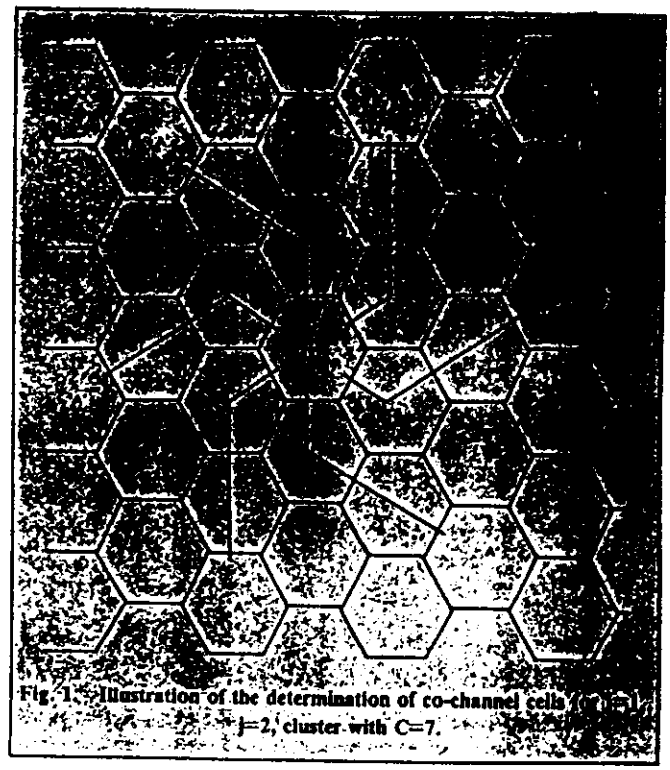


Fig. 1. Illustration of the determination of co-channel cells ($i=2$, $j=2$, cluster with $C=7$).

capacity. This frequency reuse causes co-channel interference, since the terrain or buildings in the vicinity of the mobile or the cell sites can cause a receiver to block a transmitter in a more distant cell site. In order to control this effect and to evaluate the system performance, we define the reuse distance ratio D as "the ratio of the distance between cell sites that use the same channels to the cell's radius," Fig. 1, where D as a function of co-channel interference has been evaluated [10-11]. The results which have been obtained in [11] reveal that to avoid high levels of co-channel interference it is necessary to use large D.

The cells form a natural block or cluster around the reference cell in the center and around each of its co-channel cells. The exact shape of a valid cluster is not unique, all that is required is that it contains exactly one cell with each label. Figure 1 shows a cluster with C = 7. The number of channels in each cell, $N_S = \frac{N_t}{C}$, where N_t is the total number of channels available for the system, and C is the cluster size given by:

$$C = (i + j)^2 - ij$$

Here, i and j are positive integers including zero.

Initially, a fixed number of channels N_S per cell are allocated. As the suburban areas surrounding the metropolitan centers grow, more hexagonal cells can be appended to the initial system. When the users' density in the metropolitan center increases and more channels than N_S are required, the cells are divided into smaller cells and the minimum reuse distance, D, which provide low level of interference is maintained. This process is called "cell splitting." Ideally, if N_S is of reasonable size to start with, no further allocation is needed as demand increases.

System Description

Communication to and from any mobile unit is made via the base station serving the cell in which the mobile unit is located. Each base station "cell site" is equipped with a controller that performs call set up, call supervision, mobile location [12], handoffs, and call termination. All base stations are connected (via wire lines) to and controlled by a central Mobile Telephone Switching Office (MTSO) which serves as a routing center, that is, mobile location and handoffs when the mobile moves from one cell to another, Fig. 2. A processor within each mobile conducts the signaling, radio control, and customer alerting functions.

Fig. 2. The basic structure of the AMPS system with cell sites (base station) located at the center of each cell.

Radio communication in AMPS employs frequency modulation. Transmission from mobiles to cell sites uses channel frequencies between 825 and 845 MHz, and from cell sites to the mobiles, frequencies between 870 and 890 MHz are used. Each band is then divided into narrow band (30 KHz bandwidth) channels, and communication is effected using channel-pair per cell.

Two types of radio channels are required. One type is called the set up channel. Set up channels transmit and receive only binary data messages. They are the common use channels and are monitored by mobiles which do not have an active call in progress. They are used only for initiating or setting up phone calls. The second type is called the voice channel, which provides the talking path for the customers and also handles short bursts of data that may be required for control purposes during the call.

When the mobile radio unit crosses the cell boundary the MTSO must re-route the call, and also an idle channel-pair in the new call must be found. If no channels are available in the new cell, the call is permitted to degrade until a channel is available.

Channel Assignment in Cellular System

In a cellular mobile radio system, the number of channels in each cell is determined by the user density, the frequency reuse distance D (to produce an acceptable co-channel interference), and the available bandwidth. After a base station has been assigned to serve the mobile radio unit, a channel assignment procedure must be followed to determine if a channel is available or not.

In a fixed channel assignment [13] scheme, a subset of the total available channels is permanently assigned to serve a certain cell. The channel subsets are reused in the coverage area separated by a reuse distance. Only channels from this subset can be used to serve a call within the cell. If all the channels in this subset are busy, then service cannot be provided (blocked call) even though there may be vacant channels among those which are assigned to serve in adjacent cells.

In dynamic channel assignment scheme [13-15] all channels are kept in a central pool, and any channel can be used in any coverage area "cell." In order to avoid a high level of co-channel interference, a channel can be reused simultaneously in another cell if they are separated by D. Control of a dynamic channel assignment system requires access to and processing of large quantities of data; a fast digital computer is required. Channels are assigned to serve calls based on the state of the system and in order to optimize some parameter within the system, different channel borrowing strategies have been proposed and simulated [13,16]. In Hybrid Channel Assignment, the total number of channels available for the entire system is divided into two groups. One group contains channels using the Fixed Channel Assignment scheme, the other group contains channels using the Dynamic Channel Assignment scheme. Simulation study has been carried out [17], to investigate in what ratio the channels should be divided. It was found that the optimum ratio depends on the percentage increase in the traffic density.

A channel assignment that reduces the co-channel interference, spurious, and intermodulation was proposed by Box [19]. A new channel assignment scheme has been proposed and simulated in [20]. This scheme uses a flexible fixed channel assignment with borrowing and channel reassignment in such a way as to minimize the blocking probability. The results reveal that such strategy has a better performance than other known assignments.

In the next section we shall discuss some of the diversity techniques useful in improving the communication in mobile radio.

Diversity Techniques for Land Mobile Radio

Different types of diversity, such as space diversity, frequency diversity, polarization and angle diversities, and time diversity, can be used to improve the performance of communication systems. (Diversity in land mobile radio is used to combat the effect of Rayleigh fading signal discussed earlier.)

Time diversity, as the name suggests, involves the repetition of messages and hence, delays the effective information transfer. Sequential amplitude samples of a randomly fading signal, if separated sufficiently in time, will be uncorrelated with each other and therefore, the samples provide "independent information." However, the minimum time separation between samples is inversely proportional to the speed of the vehicle. In other words, for the stationary vehicle, the time diversity is useless. Even though the delay can be tolerated, the above fundamental limitation rules out time diversity for mobile radio. Perhaps space and frequency diversity are the best applicable.

Space diversity has the advantage that it does not need additional spectrum. The basic requirement is that the spacing of the antennas in the receiving or transmitting array be chosen so that the individual signals are at least partially uncorrelated. Recalling the discussion on channel characteristics, we see that the antenna spacing of roughly one half wavelength could be sufficient. At 850 MHz, this means a fraction of a meter separation, which can be achieved in mobiles. However, for diversity at the base station, since the important scatterers are in the immediate vicinity of the mobile, the base station antennas must be considerably farther apart to achieve decorrelation. Separation of the order of tens of wavelengths would probably be adequate at the base station [21]. We shall subsequently discuss some of the combining schemes.

Instead of transmitting the desired message over spatially separated paths, as described in space diversity, one can employ different frequencies to achieve independent diversity branches. The frequencies must be separated enough so that the fading associated with different frequencies is uncorrelated. If the spacing between the carriers is sufficiently larger than the coherence bandwidth, then independent fading of the two signals can be expected. The price paid is increased spectrum usage. However, by use of spread spectrum concepts, it may be possible to achieve "frequency diversity" without the loss of spectral efficiency.

Specific Space Diversity Combining Techniques for Mobile Radio

Selection diversity—this is perhaps the simplest technique of all. Referring to Fig. 3, we see that one of the M receivers having the highest baseband SNR is connected to the output. As far as the statistics of the output signal are concerned, it is

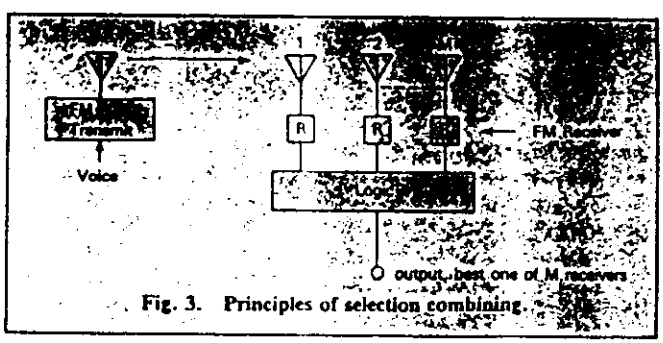


Fig. 3. Principles of selection combining.

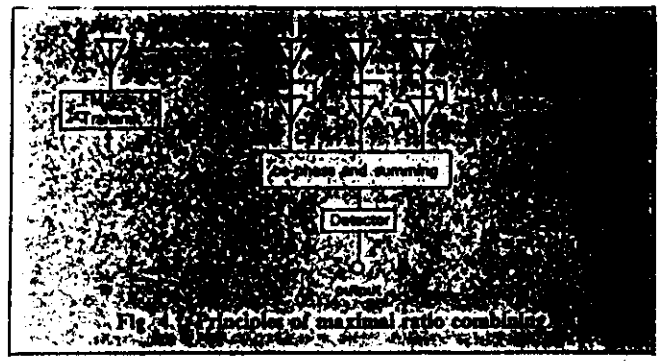


Fig. 4. Principles of maximal ratio combining.

immaterial where (*IF* or *RF* or at the antennas) the selection is done. Two branch diversity can improve the signal level by 10 dB at the 99 percent reliability level (called the diversity gain); four-branch diversity yields 16dB diversity gain.

Maximal ratio combining—here the M signals are weighted proportionally to their individual signal voltage-to-noise power ratios and then summed. Figure 4 shows the essentials of the method. The individual signals must be co-phased before combining, in contrast to selection diversity. This kind of combining gives the best statistical reduction of fading of any known linear diversity combiner. A two branch diversity gain of up to 11.5 dB gain at 99 percent reliability level can be achieved, and four branches can give 19 dB gain. It may now always be convenient or desirable to provide the variable weighting capability required for true maximal ratio combining. Instead, the gains may all be set equal to a constant value of unity, and "equal gain combining" results. This is only a fraction of a decibel poorer than maximal ratio.

The AMPS system can use space diversity at both ends (see voice and data transmission) although mobile manufacturers typically do not yet offer diversity at the mobiles. Improved systems, with many branch space diversities are under investigation [57,58].

Spectrum-Efficient Technology for Mobile Radio

With the availability of 40 MHz in the 800 MHz band for land-mobile radio, the interest focused on the effective utilization of the frequency spectrum, that is accommodating the maximum number of users in a given geometrical location, within the available bandwidth and at a reasonable cost. While three possible methods have been proposed, no clear-cut decision on any method taking into consideration all the aspects, could be made yet. In [22-23] possibility of a single side band (SSB) amplitude modulation transmission to replace the existing FM mobile radio was discussed. Though the author has predicted seven to ten times increased spectrum-utilization over FM, later reports [24] raised doubts about the suitability of SSB as an alternative. The motivation to use SSB is due to the fact that the modulated SSB signal occupies less bandwidth than a narrow band FM signal.

Another possibility is to reduce the channel spacing to 12.5 KHz in an FM system. Finally, there's the possibility of spread-spectrum modulation. At first, someone could wonder how the spreading of bandwidth occupied by a user could lead to increased spectral efficiency. In fact, it is possible to reduce the interference among the users by properly spreading the transmitted energy, thereby accommodating many users over a given bandwidth, as will be seen in the next section.

Spread Spectrum for Mobile Radio

The first attempt to introduce spread spectrum as a spectrally efficient scheme for mobile radio appeared with the proposal of

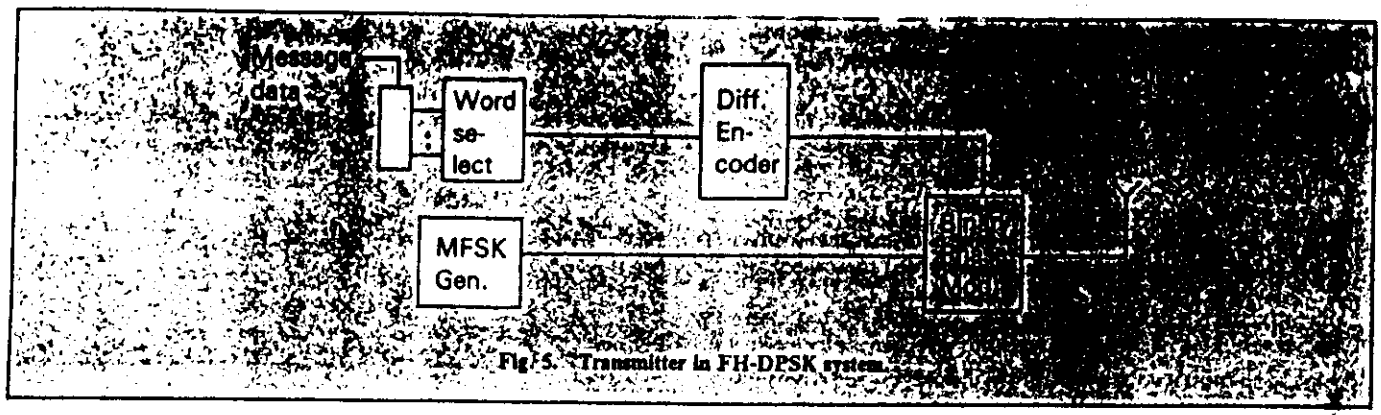


Fig. 5. Transmitter in FH-DPSK system.

a Frequency Hopped-Differential Phased Shift Keyed system (FH-DPSK) [25,27]. Generally, a direct sequence spread spectrum system performs very poorly under multi-user interference and fading channels [28]. Frequency hopping (FH) schemes possess inherent capability to provide diversity against frequency-selective fading encountered in mobile channel. Henceforth, we shall consider only the FH schemes. Some of the advantages of the spread spectrum which are not easily achievable with other systems are summarized in [25]. With spread spectrum, the questionable cost effectiveness and the formidable technological problems in the implementation remains to be answered [26].

A preliminary analysis of FH systems [27,29,30] shows that these systems could be spectrally more efficient than the conventional FM systems. Another analysis of the FM system with several branch diversities indicates that this system could be more efficient spectrally than the FH systems. However, all these studies do not take into account all the relevant mobile radio conditions, and until experimental results and further analytical work are made available, no decisive choice could be made. An alternative to FH-DPSK is the Frequency Hopped Frequency Shift Keying (FH-FSK) and this has been shown to accommodate more simultaneous users than the former scheme. In the following paragraphs, we shall discuss these two systems and summarize the available results.

FH-DPSK System [25,27]

Transmitter

A block diagram of the FH-DPSK transmitter is shown in Fig. 5. There are two parts to the modulation process in the transmitter: addressing and encoding. Addressing is performed by the MFSK generator, which repeats with period T a specific sequence of N different tones or chips, each of duration t_f ($t_f = T/N$). The specific sequence is generated according to an assigned address to the user and each user is assigned an address which is distinguishable from others despite overlap in some positions. Each mobile is fitted with a transmitter of the kind described and with a receiver to be described. The power radiated from each mobile is remotely controlled from the base station to ensure that all the signals arriving at the base station are nearly of equal strength. If this is not done, mobiles close to the base station will swamp the signals of those further away. This is the "near-far" problem common to most spread spectrum systems.

Signal information is impressed or encoded onto the MFSK address sequence in the form of binary differential phase shift keying. If a binary-1 is to be transmitted in the i^{th} chip of the address sequence, the phase of that chip is changed by π radians relative to the phase in the i^{th} chip of the previous sequence. For a binary +1, no phase change takes place. In order to increase the resistance of this type of modulation to interference, the allowed phase modulation sequences or words are selected from a set of N orthogonal words, such as the columns of Hadamard matrix [31].

Receiver

A block diagram of a typical receiver is shown in Fig. 6. There are N sections each with a band pass filter, delay element, product detector, and a low pass filter. Each section is typical of a receiver used to detect DPSK signals [32]. The array of t_f second delay lines and the set of N band pass filters selects the desired address waveform out of the incoming signal. In other words, the band pass filter center frequencies ($\omega_1 \ldots \omega_N$) are uniquely related to the address of the user under consideration. Each band pass filter is matched to rectangular chip of duration t_f and therefore, has a noise bandwidth of $1/t_f$. All N chips pass through the filters at the same time and

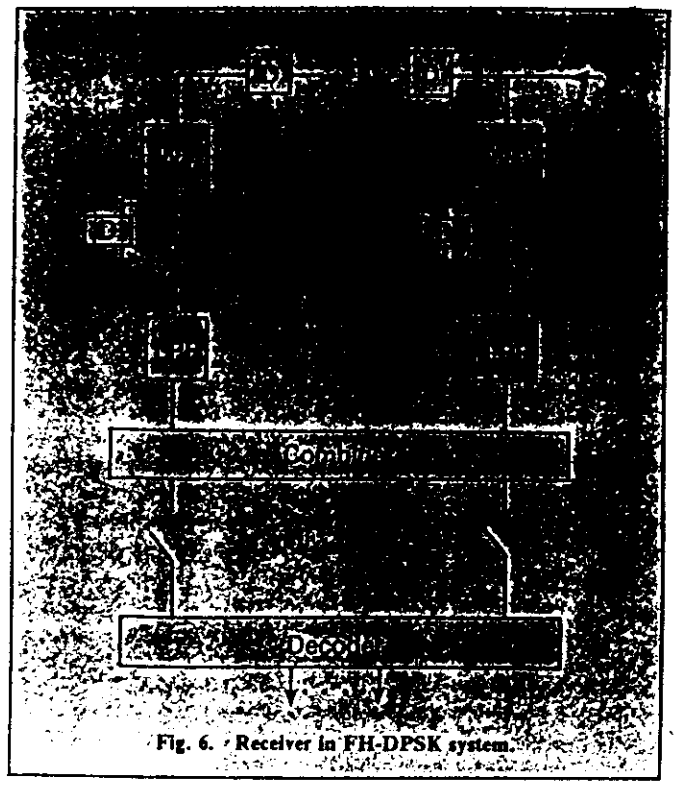
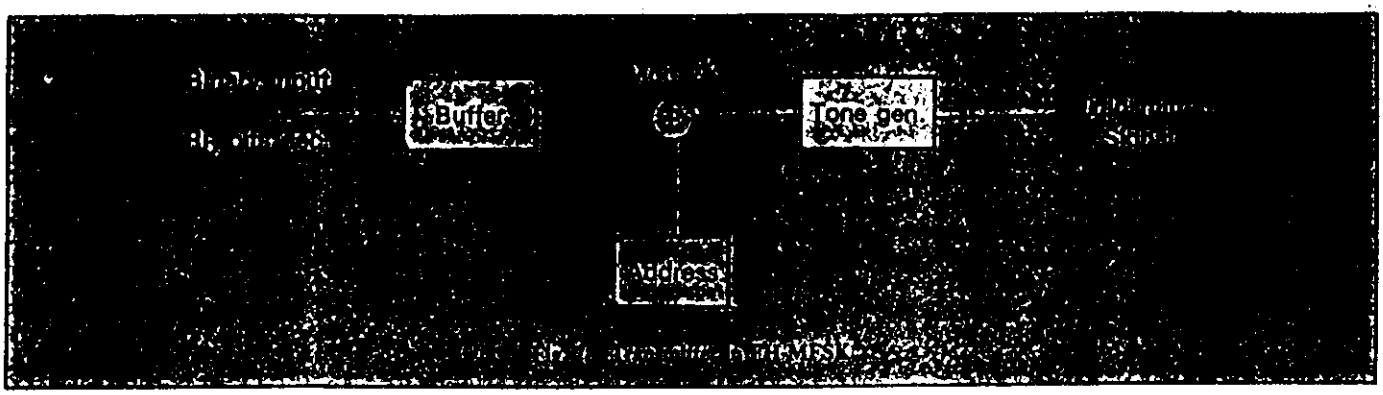


Fig. 6. Receiver in FH-DPSK system.



their phases (relative to the previous word) are detected using the T sec. delay element and the product detectors. After low pass filtering to remove the second harmonic product terms, the detector outputs are processed in a combiner circuit.

Combiner Circuits

Denoting the outputs of low pass filters as X_ℓ, $\ell = 1,\ldots N$ and the outputs of combiner as C_k, $k=1, \ldots N$ we can express C_k's as:

$$c_k = \sum_{\ell=1}^{N} h_{k\ell} X_\ell$$

where $h_{k\ell}$ are the elements of $N \times N$ Hadamard matrix. Due to channel characteristics the possibility of a word detection error arises. It turns out that the linear combiner (above) is not the best to use in order to minimize the probability of error. When $X\ell$ can be statistically characterized under Rayleigh fading and with some interference models, it is possible to apply hypothesis testing to determine the correct word. This leads to a likelihood receiver, which is the best in the sense of minimizing the probability of bit error, with the available information [33].

For successful operation of FH-DPSK under multi-user environment, we need to assign addresses that possess certain properties. Since transmission from the mobile units to the base stations are to be non-synchronous (which is advantageous), it is desirable to use signal sets that have uniformly small cross-correlation functions for any relative time shift. A class of time frequency coded signal with the one-coincidence property is suitable for the purpose [25].

Assuming ideal conditions, that is no fading, no receiver noise and perfect synchronization, Rowe has derived upper bounds on the active users U for the best possible signal set as functions of W, R_b, alphabet size, and P_b [36]. Whereas these results do not correspond to actual conditions, they predict the best that is possible under ideal situations. In [37] an upper bound on number of vectors (this when divided by number of vectors per user gives the number of users) possible on d dimensional space (here $d = 2TW$) is obtained, given some suitably defined root-mean square cross correlation C_{rms}.

Performance Analysis of FH DPSK System

The maximum number of simultaneous users that could be accommodated at a specific bit error rate is of interest in analyzing a digital mobile radio system. Since this number could vary depending on the available bandwidth or on the information bit rate, a measure defined as "spectral efficiency" is also useful:

$$\eta = \frac{M R_b}{W}$$

where η is spectrum-efficiency, M is the number of users simultaneously served by the system, R_b is the bit rate per user, and W is the one-way bandwidth.

We shall discuss later the merit of η as a measure of spectral-efficiency in a larger context of the efficient utilization of the spectrum allocated for land-mobile users. Presently, η can be used to compare two cellular systems. A comparison of average number of usable "channels" per cell for FH-DPSK and FM systems showed that they do not differ greatly. In [38], it is noted that no error occurs if the transmission is frame synchronous (which is possible for base to mobile communications), and for frame-asynchronous mobile to base communications, a good estimate of P_b was arrived at. The analysis showed that only 26 users can be accommodated at $P_b < 10^{-3}$ and at large E_b/N_o with $W = 20$ MHz, $R_b = 32$ kb/s, and orthogonal coding rate $\lambda = 5/32$ ($= \log_2 N/N$). This figure improved to $M = 46$ from 26 by the use of hard-limited combing [34]. However, likelihood combining did not give any further significant improvement [33]. Another analysis of P_b with different models of fading for the envelope of received signal was carried out in [39]. In all the above analyses, shadow fading was neglected. Also, the assumption was made that the frequency separation between the spectrum of the hopping signals was greater than the coherent bandwidth of the mobile channel. Even though this gives the full benefit of diversity and makes the analysis simpler, this may not be strictly valid [40].

Next, we shall discuss another FH scheme which is more efficient than FH-DPSK.

Frequency-Hopping Multi-Level Frequency-Shift Keyed System (FH-MFSK)

A block diagram of the m^{th} transmitter of FH-MFSK system is shown in Fig. 7. Figure 8 shows the block diagram of the receiver. The operation of the system can be understood by referring to Fig. 7, and Fig. 8. Every T seconds, K message bits are loaded serially in a buffer and transferred out as a K-bit word X_m. Assuming the modulo-2^K adder does nothing for the moment, X_m will select one of the 2^K possible different

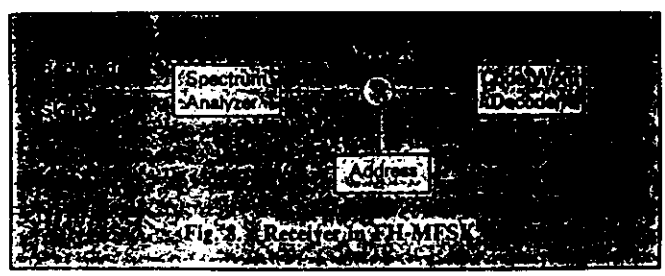


Fig. 8. Receiver in FH-MFSK

frequencies from the tone generator. At the receiver, the spectrum of each T second transmission is analyzed to determine which frequency, and hence, which K-bit word, X_n is sent. Of course, the system as such is useless for multiple-user operation. If a second transmitter were to generate X_n, neither the receiver m nor the receiver n would know whether to detect X_n or X_m. To avoid this, we add the address generator as shown in Fig. 7 and assign a unique address to each user.

The basic interval T is divided into L intervals of duration τ each. Over T seconds, the address generator of m^{th} user generates a sequence of L numbers:

$$a_m = (a_{m1}, a_{m2}, \ldots a_{mL})$$

$$Each\ a_{m\ell} \in \{0, 1, 2, \ldots 2^K - 1\}$$

whereas in [30] each $a_{m\ell}$ is selected at random from the set specified above (called "random address assignment"). There are certain addressing schemes [41,42] which possess certain algebraic structure and hence, when decoded at the receiver properly, could lead to better performance.

Now, each $a_{m\ell}$ is added modulo-2^K to X_m to produce a new K-bit number:

$$Y_{m\ell} = X_m + a_{m\ell}$$

or,

$$\underline{Y} = (Y_{m1}, Y_{m2}, \ldots Y_{mL})$$

$$\underline{X_m} = (X_m, X_m, \ldots X_m)$$

$$\underline{Y} = \underline{X_m} + \underline{a_m}$$

Each τ seconds, $Y_{m\ell}$ selects the corresponding transmitter frequency. At the receiver, demodulation and modulo 2^K subtraction by the same number $a_{m\ell}$ are performed every τ seconds, yielding:

$$Z_{m\ell} = Y_{m\ell} - a_{m\ell} = X_m$$

The sequence of operations is illustrated by the matrices of Fig. 9a and 9b. Each matrix is either a sequence of K-bit numbers (codeword, address, detection matrix) or a frequence-time spectrogram (transmit spectrum, receive spectrum). The matrices pertain to one link in a multi-user system. Crosses show numbers and frequencies generated in that link. Circles show the contributions of another link. As stated earlier, the transmit spectrum is generated by modulating the address with a code word using modulo-2^K addition. Equivalently, when each entry in the address matrix is shifted cyclically by the row number specified by the codeword matrix, we get the transmit spectrum (Fig. 9a).

Because of multi-users, extraneous entries are created in the detection matrix. For example, a word X_n transmitted over the n^{th} link will be decoded by the receiver m as:

$$Z'_{m\ell} = X_n + a_{n\ell} - a_{m\ell}$$

The $Z'_{m\ell}$ are scattered over different rows. The desired transmission, on the other hand, is readily identified because it produces a complete row of entries in the detection matrix. Normally, the fading of the tones and the receiver noise can cause a tone to be detected when none has been transmitted (false alarm) and/or can cause a transmitted tone to be undetected (miss). Even without these impairments, many user entries can combine to produce a complete row other than X_m and hence, can cause errors in the identification of a correct row (in the word X_m). Hence, a majority logic rule is attempted: select the codeword associated with the row containing the greatest number of entries. Under this decision rule, an error will occur when insertions (detected tones due to other users and false alarms) combine to form a row with more entries than the row corresponding to the transmitted code

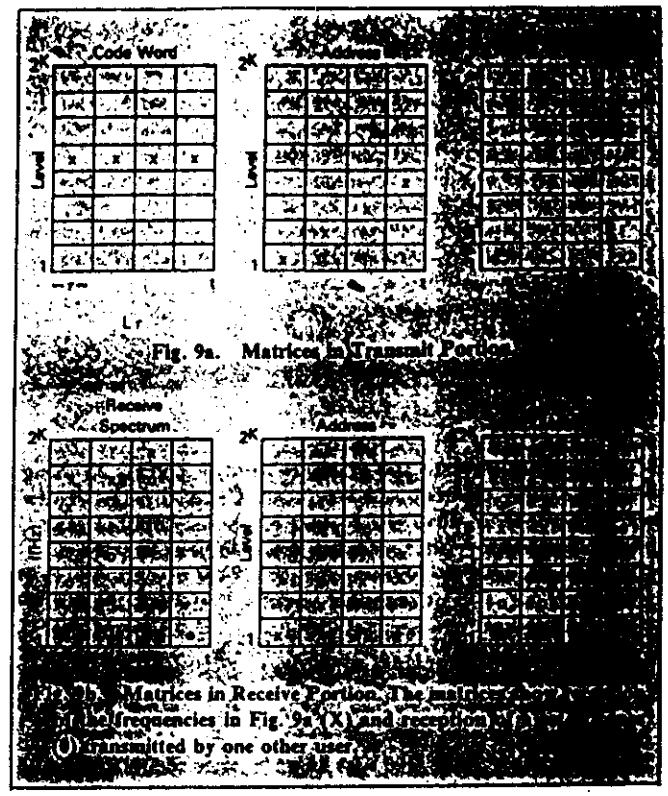


Fig. 9a. Matrices in Transmit Portion

Fig. 9b. Matrices in Receive Portion. The matrix ... frequencies in Fig. 9a (X) and reception ... transmitted by one other user.

word. An error can occur when insertions combine to form a row containing the same number of entries as the row corresponding to the transmitted code word. We view the transmission to each square in the tone detection matrix as an example of non-coherent, on-off keying. Because of fading of amplitude and the random change of phase, it is not possible to employ coherent detection in a mobile environment. (Recall that we used differential phase detection in FH-DPSK scheme, since the phase is not likely to change significantly from bit to bit [25].) From the textbook formulas we have [43]:

$$P_F = \exp(-\beta^2/2)$$

$$P_D = 1 - \exp(-\beta^2/2(1+\bar{\rho}))$$

where P_F denotes false alarm probability, P_D the probability of deletion (miss), β the normalized threshold set in the receiver and $\bar{\rho}$ the average signal to noise ratio. The above scheme, where the presence or absence of energy in each square of detection matrix is decided, together with majority logic decision is called "Hard-limiting Combining" [44]. Whereas the hard-limited receiver does not fully exploit the available statistics of the received signal, it can be shown to be not too inferior to an ideal likelihood receiver [45,33].

The results on spectrum efficiency η, defined earlier, show that η is maximum for W = 20 and decreases slightly as W is reduced to 5 MHz [46]. Thus, splitting the total available bandwidth of 20 MHz into smaller bandwidths will only lead to reduced efficiency. The effect of shadowing on the performance of hard-limited receiver was examined in [47]. The results show that the system capacity decreases to 130 users at $P_b < 10^{-3}$ for average SNR = 30 dB and log-normal shadowing standard deviation of 12 dB. Also, in cellular systems, additional degradation occurs due to interference from adjacent cells. It could be seen that the performance deteriorates rapidly without power control. A power control scheme was suggested and evaluated for base to mobile communication [48]. It was observed that with Rayleigh fading and an SNR of 25 dB, the number of users that could be accommodated in a single cell at

$P_b < 10^{-3}$ was about 115. This is a reduction of 55 users from the isolated cell case.

In the above analysis, it was assumed that the frequency separation between tones, namely $1/\tau$, was greater than the coherence bandwidth of the channel and hence, the assumption that all the tones fade independently. For typical $\tau = 11.6\ \mu$ sec., $1/\tau = 86$ KHz. However, the coherence bandwidth (defined with respect to correlation coefficient of 0.9) varies from < 0.04 MHz in urban areas to > 0.25 MHz, < 1 MHz in suburban areas [49]. Thus, it is possible that correlated fading could occur during deep fades and lead to error clustering [50,51].

Improved Address Assignment and Decoding

In FH-MFSK system considered so far, we assumed that the addresses are assigned randomly. Also, the probability of bit error P_b computed is the value obtained by averaging all the probability of bit errors resulting from an ensemble of all possible random address assignments. Naturally, this raises the question, "are there bound to be many links whose P_b will be very high, even though the average $P_b < 10^{-3}$?" We find the answer [52], "since only a fraction $1/\lambda$ of a set of positive numbers can be larger than λ times their average, at least 90 percent of all codes in the collection must have a P_b no greater than ten times average P_b, and 99 percent of all code must have a P_b no larger than 100 times average P_b." In fact, the result of this argument is verified by simulation studies [53]. The simulation is done assuming perfect channel, but multi-user interference is considered. It is noticed that maximum BERs (bit error rates) do not differ by more than a factor of 2 from the average BER. However, if the addresses are assigned in a "best possible way," each user could have a nearly identical performance. One possible scheme is discussed next.

In [41], the proposed addressing scheme assigns each user an address

$$\underline{a_m} = (\gamma_m, \gamma_m\beta, \gamma_m\beta, \ldots \gamma\beta^{L-1})$$

where $\gamma_m \in GF(Q)$

β is any primitive element of GF(Q).

Galois Field exists only for Q being prime or any integer power of a prime (Hence, Q can be 2^K) [35]. Coding of the user message m is done according to:

$$\underline{Y_m} = \underline{a_m} + \underline{x_m}$$

Observe that this equation is similar to a previous equation. It is easily shown that for synchronous communication, any two coded messages $\underline{Y_m}$ and $\underline{Y_e}$ ($m,e \in GF(Q)$, $m \neq e$) could coincide in one chip, at most. Then, it is clear that under ideal transmission, to create one spurious row in the decoded matrix at a user, at lease L users must have been involved. An error can occur in the decision process only if one or more spurious rows are created. A simple upper bound on probability of bit error indicates that this bound is slightly less than the average P_b bound obtained for random addressing. This agrees with our expectations.

Since the addressing scheme (above) possesses certain algebraic properties, it is possible to exploit these properties in the decoding of messages, thereby accommodating more number of users at a specified P_b, (say $< 10^{-3}$). Such an analysis carried out in [54] under ideal conditions shows that the scheme could accommodate nearly 450 users as compared to 216 with conventional decoding. Of course, the decoding scheme requires the knowledge of the addresses of all active users. Therefore, such information needs to be periodically transmitted to mobiles from base. Another difficulty is that the decoding procedure does not allow an easy performance evaluation under the conditions of noise and fading. Also, the complexity of decoding at the mobiles could be very high, making the algorithm useless under such a situation.

It is also possible to use space diversity of moderate complexity (2 or 3 branches) to improve the performance of FH-MFSK [4]. There are some recent results suggesting the possibility of space diversity with FM or PSK [55]. However, the number of diversity branches needed is greater than 20 or so and hence, the implementation becomes formidable. Another version of space diversity with PSK and time-division retransmission is suggested in [56].

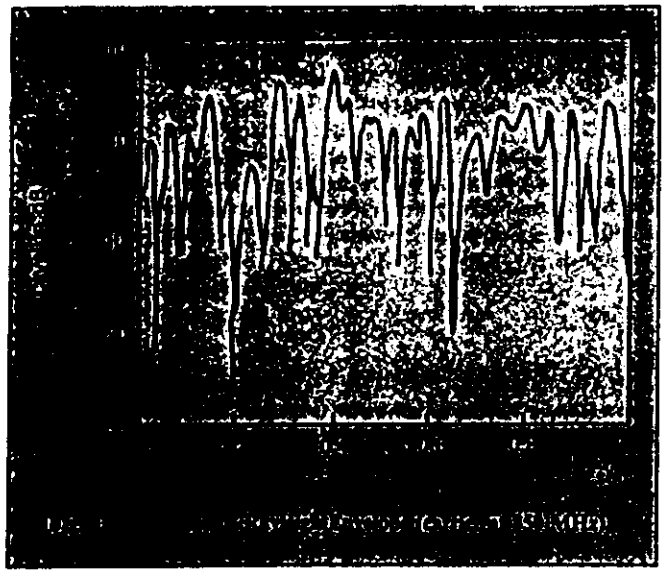
Measures for Spectral Efficiency

In previous sections we used η ($= \frac{MR_b}{W}$) as a figure of spectral efficiency. Apart from other drawbacks of this measure, there is one which has not been observed. With the use of efficient speech coding techniques, such as LPC, it is possible to reduce R_b to as low as 4 Kb/sec. Assume such coded speech, when passed through mobile radio channels, still possesses intelligibility. Then, with FH-systems it can be shown that M would increase considerably. However, because of the multiplication of M by R_b in the definition of η, the net increase (it could be a decrease too!) may not be much, and as such this is clearly misleading because LPC speech also carries effectively the same information over the same period. Hence, η is code dependent and cannot be applied to compare two systems employing different coding techniques. Some of the measures of spectral efficiency for land-mobile application and their usefulness are discussed in [57]. In general, a definition should include the number of mobiles served over what bandwidth and geographical area. Clearly η, as defined earlier, does not include the geometry factor. Finally, a measure which is useful for FM cellular systems expressed in Erlangs / H_z m^2. The reader is cautioned that assigning a minimum bandwidth per channel does not always lead to an overall spectrally efficient FM system. For example, allocation of 30 KHz per channel can be spectrally more efficient than an assignment of 15 KHz per channel [58]. Next, we discuss some of the aspects of voice and data transmission in FM-cellular systems, with specific references to AMPS.

Voice and Data Transmission

Voice Transmission

Speech can be transmitted by an analog modulation, such as FM, or by means of a digital radio, after coding the speech. While AM is conceptually possible, the rate of change and the depth of fades that can occur at UHF have not permitted satisfactory transmission quality to be obtained in this environment. We shall explain briefly the impairments expected on voice transmission and the measures possible to reduce these, when FM is employed. Specifically, we concentrate on AMPS (Advanced Mobile Phone System) [59,60]. Considering 850 MHz band, Fig. 10 shows a sample of Rayleigh envelope signal obtained at moving antenna measured along a short distance of travel. We observe that the Rayleigh fades occur approximately one-half wavelength apart. At carrier frequencies near 850 MHz, independent fades are about seven inches apart. As the mobile receiver moves through the radio interference pattern, it is therefore subjected to frequent fades. Figure 11 shows the probability distribution function of the received instantaneous signal power normalized to its mean-value. The no-diversity curve shows that the fades are such that 10 percent of the time the signal will be 10 dB below its local mean, 1 percent of the time 20 dB will be below the mean, and so forth. Also of interest are the quantities, the level crossing rate of the envelope below a specified level, as well as the duration of the fade below the specified level. These two quantities are



dependent on the vehicle speed. Apart from this Rayleigh fading, we have the "shadow fading" caused by terrain features.

In AMPS, discriminator detection is employed to demodulate voice signals. During the deep fades, the signal goes below noise level and the noise could "capture" FM receiver. These interruptions have a different subjective effect as a function of the speed. This is also called the "click noise." These clicks arrive in bursts and are time correlated with RF signal fades. At times during fades, the receiver may be captured by interference from a co-channel user. The result is a burst of interfering voice modulation which is unintelligible because of the short duration of fades. Sometimes, the frequency offset between the signal carrier and the co-channel carrier can be heard as a wobbling tone. Apart from these, we have the impulsive noise (due to ignition systems) which is predominant in urban areas. Since majority of the impairments come through fading, it is important to reduce the level, duration, and frequency of fades. This is accomplished by "diversity techniques" discussed earlier. Space diversity is attractive with FM systems. The effect of a two-branch equal gain combining diversity system is shown in Fig. 11. In AMPS, equal gain or selection diversity will be used at cell sites, whereas the mobiles, where the cost and complexity are important, will be provided with switched diversity. However, performance of switched diversity is not as good as equal gain combining [3]. In addition to the improvement through diversity, additional improvement with voice processing circuits are necessary and possible. For example, variations in talker volume can have significant effects on the subjective quality of the received signal. Low volume speakers induce low frequency deviations and hence, the received signal will be weaker. In contrast, speech from loud talkers is impaired through excessive clipping distortion in the transmitter. To overcome these problems, syllabic companders at the transmitter and the expanders at the receivers are employed.

Digital coding and transmission of speech has some attractive features, namely, inexpensive coder-decoder implementation, straight forward speech encryption (bit scrambling), and efficient signal regeneration. But in a mobile radio environment, the digital transmission too, faces the problem of fading. Without any effective diversity, errors tend to occur in bursts. Whereas an average bit error rate of 1 in 10^3 still gives good quality for ADM speech samples, an error rate of 1 in 10^2 is acceptable during short periods. However, since these errors do not occur independently but rather in bursts, they have annoying effects on the listener. There has not been much discussion on this subject, except for a few papers. In [61], a differential code with explicit transmission of step size over an error protected channel, was analyzed for its performance by simulation studies. Effect of redundant time diversity coding and bit scrambling was also investigated.

Data Transmission

In the AMPS system discussed above, direct binary frequency shift keying of the carrier with discriminator detection was employed. The main source of impairment is once again the error clustering during fade. Use of error correcting codes (40, 28 BCH code) along with message repetition and majority voting at the receiver are considered as measures against burst errors.

Recently, there has been interest in finding efficient modulation techniques for data transmission in telephone, as well as, satellite networks [62,63]. In telephone lines, the bandwidth is at a premium and hence, spectrally efficient techniques to send data at higher rates are important. Some of these schemes, which possess narrow spectrum, also have constant envelope. The constant envelope property is important in satellite channels where the nonlinear operations used (hard-limited or class C power amplifiers) to create adjacent channel interference, if the envelope is not constantly maintained. These features are attractive for data transmission in cellular mobile radio [64]. Two modulation schemes known as MSK (minimum shift-keying), which is also called Fast FSK, and TFM (Tamed Frequency Modulation) can be considered [63,65]. MSK can be thought of as a special case of offset QPSK. In normal QPSK, the incoming data stream is split into "odd" and "even" streams and are used to modulate the quadrature and in-phase carriers, respectively, (see Fig. 12). Because of the orthogonality of the two carriers, it is possible to recover the "odd" and "even"

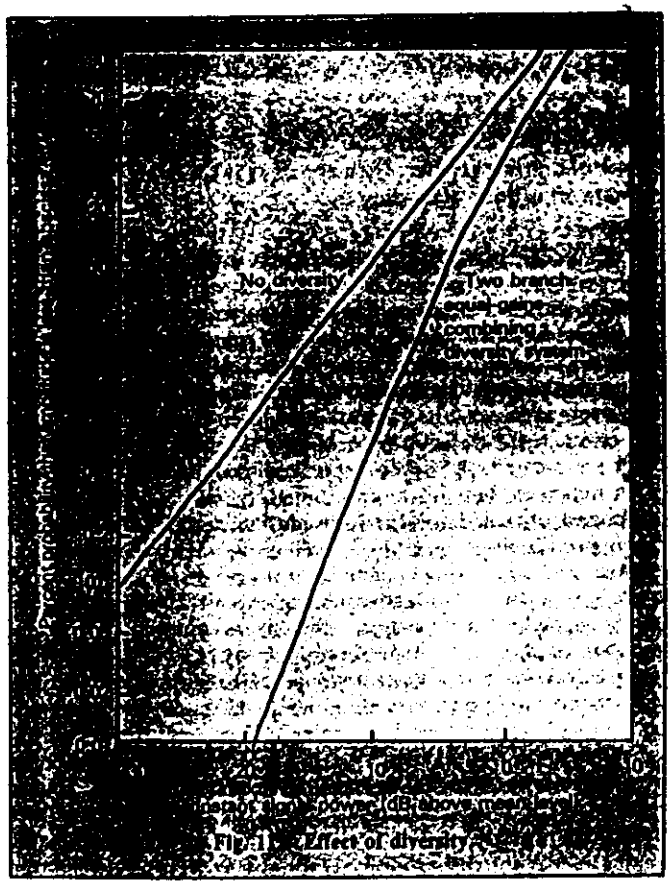


Fig. 11. Effect of diversity

streams exactly at the receiver, when there is no channel impairment. Whereas in QPSK, the bit transitions occur at the same time, in OQPSK, their occurrences differ by $\pi/2$ radians.

Because of this, the phase changes in OQPSK waveform at bit transitions can only be $\pm\pi/2$, which is in contrast to $\pm\pi$ and $\pm\pi/2$ changes possible in QPSK. This helps in out of band radiations to be reduced in OQPSK. With sinusoidal pulse weighting in offset DPSK the resulting waveform S(t) can be written as (Fig. 12):

$$S(t) = a_I(t)\cos\left(\frac{\pi t}{2T}\right)\cos(2\pi f_c t) + a_Q(t)\sin\left(\frac{\pi t}{2T}\right)\sin 2\pi f_c t$$

Using standard trigonometric formulae this can be rewritten as:

$$S(t) = \cos\left[2\pi f_c t + b_k(t)\frac{\pi t}{2T} + \phi_k\right]$$

Where $b_k(t) = -a_I(t)\,a_Q(t)$ and ϕ_k is 0 or π corresponding to $a_I = 1$ or -1.

From the above equation we observe the following properties of MSK:

1) It has constant envelope.
2) There is phase continuity in the RF carrier at the bit transitions.
3) The signal is an FSK signal, with continuous phase and with the frequency spacing $\Delta = (f_c + 1/4T) - (f_c - 1/4T) = 1/2T$. This is the minimum frequency spacing which allows the two FSK signals to be coherently orthogonal, hence the name minimum shift keying.

Whereas the phase is continuous in MSK, the derivative of the phase is still discontinuous (implying phase is only piecewise continuous). If the phase is made still smoother, a much narrower spectrum can be achieved. One method of modulation achieving this is called the TFM [64]. Some comparison between MSK and TFM is as follows:

1) Both are spectrally efficient, though the sharper roll-off of the spectrum of TFM can be advantageous in reducing adjacent channel interference.
2) Both are easy to generate.
3) In principle, coherent, differential, and discriminator detection can be employed for MSK and only the first two types for TFM. It can be shown that error rate performance of coherently decoded MSK is equivalent to that of PSK (Quartennary PSK) whereas, coherent detection of TFM has a loss of 1 dB. Differential detection of MSK is simple to implement but is slightly less efficient than coherent detection. A coherently detected TFM for digital speech, which performs nearly as good as conventional narrow band FM is given in [64].

In general, it is not easy to construct a simple carrier recovery circuit which enables one to regenerate the reference carrier precisely and stably in the fast Rayleigh fading channels. The performances of the differential detection receivers in mobile radio channels for MSK and TFM can be found in [66] and [67].

Clearly, some of the spectrally efficient techniques could soon find a place in the digital data transmission over mobile radio.

Conclusion

This tutorial paper looks at some of the theoretical and design issues involved in mobile radio communication. While the review is not exhaustive, many references are given to supplement the material presented.

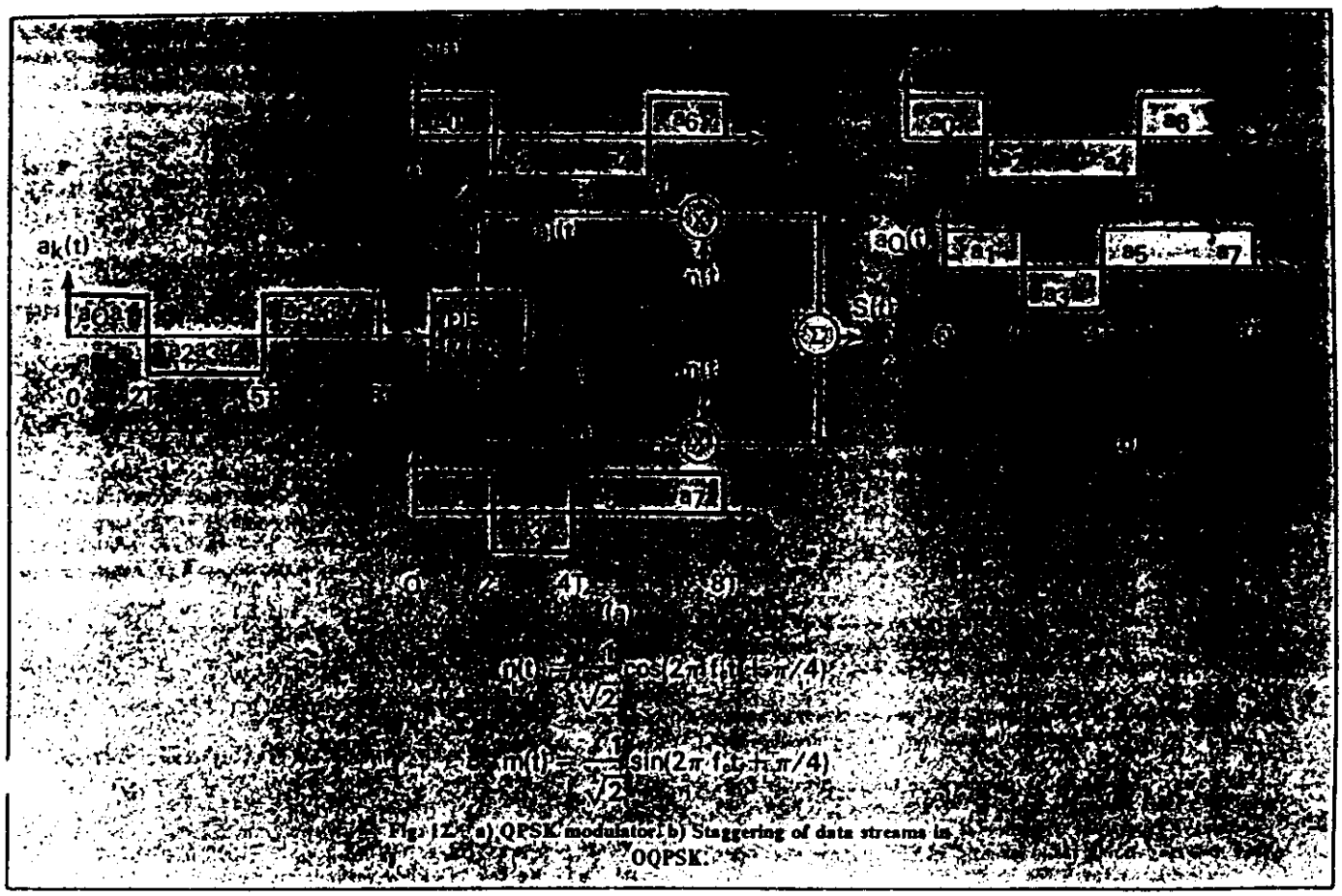


Fig. 12. a) QPSK modulator. b) Staggering of data streams in OQPSK.

References

[1] R. H. Clarke, "A statistical theory of mobile radio reception," *Bell System Tech. J.*, 47, pp. 957-1000, Jul. 1968.

[2] M. J. Gans, "A power-spectral theory of propagation in the mobile radio environment," *IEEE Trans. Veh. Tech.*, VT-21, pp. 27-38, Feb. 1972.

[3] W. C. Jakes, *Microwave Mobile Communication*, John Wiley and Sons, New York, 1974.

[4] R. Muammar, "Frequency-hopped multilevel frequency shift keying spread spectrum for mobile radio communication systems," Ph.D. Dissertation, Southern Methodist University, Dallas, 1982.

[5] D. O. Reudink, "Properties of mobile radio propagation above 400 MHz," *IEEE Trans. Veh. Tech.*, Vol. VT-23, No. 4., pp. 143-159, Nov. 1974.

[6] E. B. Glover and Dr. W. C. Adams, "Coherence bandwidth considerations for mobile radio in urban environments," *N.T.C.*, New Orleans, Nov. 1981.

[7] V. H. MacDonald, "The cellular concept," *B.S.T.J.*, Vol. 58, No. 1, pp. 15-41, Jan. 1979.

[8] W. R. Young, "Advanced mobile phone service: introduction, background, and objectives," *B.S.T.J.*, Vol. 58, No. 1, pp. 1-15, Jan. 1979.

[9] Nathan Ehrlich, "The advanced mobile phone service," *IEEE Communications Magazine*, Vol. 17, No. 2, pp. 9-15, Mar. 1979.

[10] Richard C. French, "The effect of fading and shadowing on channel reuse in mobile radio," *IEEE Trans. Veh. Tech.*, Vol. VT-28, No. 3, pp. 171-181, Aug. 1979.

[11] R. Muammar and S. C. Gupta, "Co-channel interference in high capacity mobile radio systems," *IEEE Trans. on Communication Systems*, Vol. COM-30, pp. 1973-1978, Aug. 1982.

[12] G. D. Ott, "Vehicle location in cellular mobile radio systems," *IEEE Trans. on Veh. Tech.*, Vol. VT-26, No. 1, pp. 43-46, February 1977.

[13] W. C. Y. Lee, *Mobile Communication Engineering*, McGraw-Hill, New York, 1982.

[14] D. C. Cox and D. O. Reudink, "Dynamic channel assignment in high capacity mobile communication system," *B.S.T.J.* Vol. 50, pp. 1833-1857, Jul. 1971.

[15] D. C. Cox and D. O. Reudink, "Dynamic channel assignment in two-dimensional large scale mobile radio systems," *B.S.T.J.*, Vol. 51, No. 9, pp. 1611-1629, Sept. 1972.

[16] L. G. Anderson, "A simulation study of some dynamic channel assignment algorithms in a high capacity mobile telecommunications systems," *IEEE Trans. Veh. Tech.*, Vol. VT-24, No. 4, pp. 210-217, Nov. 1973.

[17] T. J. Kahwa and N. D. Georganes, "A hybrid channel assignment scheme in large-scale cellular mobile communication systems," *IEEE Trans. on Com.*, Vol. COM-26, No. 4, pp. 432-438, Apr. 1978.

[18] K. Watanaba, "Radio channel assignment methods for mobile telephone systems," *Electronics and Comm. in Japan*, Vol. 59, No. 3, pp. 96-106, Mar. 1976.

[19] F. Box, "A Heuristic technique for assigning frequencies to mobile radio nets," *IEEE Trans. Veh. Tech.*, Vol. VT-22, No. 2, pp. 57-64, May 1978.

[20] S. Elnoubi, R. Singh, and S. C. Gupta, "A new channel assignment scheme in land mobile radio communication," Presented at NTC '79 in Washington, D.C., Dec. 1979.

[21] W. C. Jakes Jr., "A comparison of specific space diversity techniques for reduction of fast-fading in UHF mobile radio systems," *IEEE Trans. on Veh. Tech.*, Vol. VT-20, No. 6, Nov. 1971.

[22] "Spectrum efficient technology for voice communications," UHF Task Force Report, FCC, 1978.

[23] B. Lusignan, "Single-sideband transmission for land mobile radio," *IEEE Spectrum*, pp. 33-37, Jul. 1978.

[24] P. S. Henry, "Single-sideband mobile radio—a review and update; includes "Comments on UHF task force report by ad hoc committee communications divisions, for spectrum-efficient technology," *IEEE Communications Magazine*, pp. 25-27, Mar. 1979. See also pp. 2-3, Sept. 1979.

[25] G. R. Cooper and R. W. Nettleton, "A spread-spectrum technique for high capacity mobile communications," *IEEE Trans. on Veh. Tech.*, Vol. VT-27, No. 4, Nov. 1978.

[26] J. J. Mikulsky, "Mobile radio: to spread or not to spread?" *IEEE Communications Magazine*, p. 2, Jul. 1979. See also p. 2, Sept. 1979.

[27] P. S. Henry, "Spectrum efficiency of a frequency hopped DPSK spread spectrum mobile radio system," *IEEE Trans. on Veh. Tech.*, Vol. VT-28, No. 4, Nov. 1979.

[28] C. S. Gardner and J. A. Orr, "Fading effects on the performance of a spread spectrum multiple access communication system," *IEEE Trans. on Comm.*, Vol. COM-27, No. 1, Jan. 1979.

[29] G. R. Cooper, R. W. Nettleton, and D. P. Grybos, "Cellular land-mobile radio: why spread spectrum?" *IEEE Communications Magazine*, pp. 17-23, Mar. 1979.

[30] D. J. Goodman, et al., "Frequency-hopped multilevel FSK for mobile radio," *BSTJ*, Vol. 59, No. 7, pp. 1257-1275, Sept. 1980.

[31] A. J. Viterbi, *Principles of Coherent Communication*, McGraw-Hill, New York, p. 232, 1966.

[32] R. W. Lucky, J. Salz, and E. J. Weldon, Jr., *Principles of Data Communications*, McGraw-Hill, New York, p. 255, 1968.

[33] O. C. Yue, "Maximum likelihood combining for non-coherent and differentially coherent frequency-hopping multiple-access systems," *IEEE Transactions*, Vol. IT-28, No. 4, pp. 631-639, Jul. 1982.

[34] O. C. Yue, "Hard-limited versus linear combining for frequency-hopped multiple access systems in a Rayleigh fading environment," *IEEE Transactions on Veh. Tech.*, Vol. BT-30, No. 1, pp. 10-14, Feb. 1981.

[35] E. R. Berekamp, *Algebraic Coding Theory*, McGraw-Hill, New York, 1968.

[36] H. E. Rowe, "Bounds on the number of users in spread spectrum systems," National Telecommunication Conference, pp. 69.6.1-69.6.6, 1980.

[37] H. E. Rowe, "Bounds on the number of signals with restricted cross correlation," *IEEE Transactions on Communications*, Vol. COM-30, No. 5, pp. 966-974, May 1982.

[38] O. C. Yue, "Frequency hopping, multiple-access, phase shift keying system performance in a Rayleigh fading environment," *BSTJ*, Vol. 59, No. 6, pp. 861-879, Jul.-Aug. 1980.

[39] O. C. Yue, "Useful bounds on the performance of a spread spectrum mobile communication system in various fading environments," *IEEE Trans. on Comm.*, Vol. COM-28, No. 10, pp. 1819-1823, Oct. 1980.

[40] G. R. Cooper and R. W. Nettleton, "Spectral efficiency in cellular land-mobile communications: a spread spectrum approach," *Report TR-EE*, 78-44, School of Electrical Engineering, Purdue University, West Lafayette, Indiana 47907.

[41] G. Einarsson, "Address-assignment for a time-frequency-coded, spread spectrum system," *BSTJ*, Vol. 59, No. 7, pp. 1241-1255, Sept. 1980.

[42] L. I. Bluestein and R. L. Greenspan, "Efficient approximation of orthogonal waveforms," MIT Lincoln Laboratory Group Report, 1964-48.

[43] M. Schwartz, W. R. Bennett, and S. Stein, *Communication Systems and Techniques*, McGraw-Hill, New York, pp. 398-403, 1966.

[44] O. C. Yue, "Performance of frequency-hopping multiple access multilevel FSK systems with hard-limited and linear combining," *IEEE Transactions on Comm.*, Vol. COM-29, No. 11, pp. 1687-1694, Nov. 1981.

[45] R. Viswanathan and S. C. Gupta, "Performance comparison of hard-limited likelihood and linear combining receivers for FH-MFSK mobile radio," *IEEE Trans. on Comm.*, Vol. COM-31, pp. 670-677, May 1983.

[46] R. H. Muammar and S. C. Gupta, "Spectrum efficiency of a frequency hopped MFSK spread spectrum mobile radio system," *IEEE Transactions on Veh. Tech.*, Vol. VT-31, pp. 66-69, May 1982.

[47] R. Muammar and S. C. Gupta, "Performance of a frequency hopped multi-level FSK spread spectrum receiver in a Rayleigh fading and lognormal shadowing channel," International Communications Conference, *ICC '82 Conference Record*, Vol. 3, pp. 6B.6.1-6B.6.5.

[48] R. Viswanathan and S. C. Gupta, "Adjacent cell interference in FH-MFSK cellular mobile radio systems," *IEEE Trans. on Veh. Tech.*, Vol. VT-32, pp. 191-200, May 1983.

[49] G. A. Arredondo, J. C. Feggeler, and J. I. Smith, "Voice and data transmission," *BSTJ*, Vol. 58, No. 1, pp. 97-122, Jan. 1979.

[50] R. Viswanathan and S. C. Gupta, "Burst error analysis of frequency-hopping systems for mobile communication," International Communication Conference, *ICC '82 Conference Record*, Vol. 3, pp. 5B.4.1-5B.4.4.

[51] M. Mizuno, "Randomization effect of errors by means of frequency hopping techniques in a fading channel," *IEEE Trans. on Commun.*, Vol. COM-30, No. 5, pp. 1052-1056, May 1982.

[52] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*, Wiley, New York, pp. 299-305, 1965.

[53] B. G. Haskell, "Computer simulation results on frequency hopped MFSK mobile radio—noiseless case," *IEEE Trans. on Commun.*,

Vol. COM-29, No. 2, pp. 125-132, Feb. 1981.
[54] U. Timor, "Multi-stage decoding of frequency-hopped FSK system," *BSTJ*, Vol. 60, No. 4, pp. 471-483, Apr. 1981.
[55] Y. S. Yeh and D. O. Reudink, "Efficient spectrum utilization for mobile radio systems using space diversity," *IEEE Trans. on Commun.*, Vol. COM-30, No. 3, pp. 447-455, Mar. 1982.
[56] P. S. Henry and B. S. Glance, "A new approach to high capacity digital mobile radio," *BSTJ*, Vol. 60, No. 8, pp. 1891-1904, Oct. 1981.
[57] D. N. Hatfield, "Measures of spectral efficiency in land-mobile radio," *IEEE Trans. on Electromagnetic Compatibility*, Vol. EMC-19, No. 3, pp. 266-268, August 1977.
[58] D. Muilwijk, "Spectrum efficiency in mobile radio communication," *Philips Telecommunication Review*, Vol. 36, No. 1, pp. 53-60, Apr. 1978.
[59] G. A. Arredendo, J. C. Feggeler, and J. I. Smith, "Voice and data transmission," *BSTJ*, Vol. 58, No. 1, pp. 97-123, Jan. 1979.
[60] G. A. Arredendo and J. I. Smith, "Voice and data transmission in mobile radio channel at 850 MHz," *IEEE Trans. Veh. Tech.*, Vol., VT-26, No. 1, pp. 88-93, Feb. 1977.
[61] N. S. Jayant, "Step-size transmitting differential coders for mobile telephony," *BSTJ*, Vol. 54, No. 9, pp. 1557-1581, Nov. 1975.
[62] S. Pasupathy, "Correlative coding: a bandwidth efficient signaling scheme," *IEEE Communications Magazine*, pp. 4-11, Jul. 1977.
[63] S. Pasupathy, "Minimum shift keying: a spectrally efficient modulation," *IEEE Communications Magazine*, pp. 14-22, Jul. 1979.
[64] D. Muilwijk, "Tamed frequency modulation—a bandwidth saving digital modulation method, suited for mobile radio," *Philips Telecommunication Review*, Vol. 37, No. 1, pp. 35-49, Mar. 1979.
[65] R. Debuda, "Coherent demodulation of frequency shift keying with low deviation ratio," *IEEE Trans. on Commun.*, pp. 429-435, Jun. 1972.
[66] K. Hirade, et al., "Error rate performance of digital FM with differential detection in land mobile radio channels," *IEEE Trans. Veh. Tech.*, Vol. VT-28, No. 3, pp. 204-212, Aug. 1979.
[67] S. Elnoubi and S. C. Gupta, "Error rate performance of noncoherent detection of duobinary coded MSK and TFM in mobile radio communication systems," *IEEE Trans. on Veh. Tech.*, VT-30, No. 2, pp. 62-76, May 1981.
[68] R. Viswanathan and S. C. Gupta, "Adjacent cell interference in FH-MFSK cellular mobile radio system," *IEEE Trans. Veh. Tech.*, Vol. VT-32, pp. 191-200, May 1983.
[69] R. Viswanathan and S. C. Gupta, "Non parametric receiver for FH-MFSK mobile radio," *IEEE Trans. on Comm.*, Vol. COM-33, pp. 178-184, Feb. 1985.
[70] John Oetting, "Cellular mobile radio—an emerging technology," *IEEE Communications Magazine*, Vol. 21, No. 8, pp. 10-15, Nov. 1983.

Someshwar C. Gupta received the B.A. (Hons.) and M.A. degrees in Mathematics, in 1951 and 1953, respectively, from Punjab University, India, the B.S. (Hons.) degree in Electrical Engineering, from Glasgow University, Scotland, in 1957, and the M.S.E.E. and Ph.D. degrees in Electrical Engineering from the University of California, Berkeley, in 1962.

He has considerable industrial and teaching experience, and is currently Professor and Chairman of the Department of Electrical Engineering, Southern Methodist University, Dallas, Texas. He is also a member of the Technical Review Committee for Western Union Company. Dr. Gupta is author of the book *Transform and State Variable Methods in Linear Systems* (Wiley, 1966) and coauthor of the books *Fundamentals of Automatic Control* (Wiley, 1970) and *Circuit Analysis—With Computer Applications to Problem Solving* (Matrix, 1972).

He was the Consulting Editor for International Textbook's series on *Circuits, Systems, Communications and Computers.* He is currently a member of the Editorial Board of *International Journal of Systems Sciences.*

R. Viswanathan received the B.E. (Hons.) degree in electronics and communication engineering from the University of Madras in 1975 and the M.E. degree with distinction in Electrical Communication Engineering from the Indian Institute of Science, Bangalore, India, in 1977. He received the Ph.D. degree in Electrical Engineering from Southern Methodist University, Dallas, Texas in May, 1983.

From 1977 until 1980 he worked as a Deputy Engineer in the Digital Communication Department of the Research and Development Division of Bharat Electronics Ltd., Bangalore, India. Presently, he is an Assistant Professor of Electrical Engineering at Southern Illinois University, Carbondale, Illinois. His research interests include mobile radio communication and spread spectrum techniques. He is a member of Tau Beta Pi.

Rushdi H. Muammar received the B.S. (second class honors) degree in Electrical Engineering from Baghdad University, Iraq, in 1974, and the M.S.E.E. and Ph.D. degrees Southern Methodist University, Dallas, Texas in 1979 and 1982 respectively.

From 1974 to 1978 he worked with the Major Project Department, Kuwait. Since 1982 he is an Assistant Professor in Electrical Engineering at the University of Puerto Rico, Mayaguez. His current field of research encompasses problems related to mobile radio communication. ■

Universal Digital Portable Communications: A System Perspective

DONALD C. COX, FELLOW, IEEE, HAMILTON W. ARNOLD, SENIOR MEMBER, IEEE, AND PHILIP T. PORTER, SENIOR MEMBER, IEEE

Abstract—In our highly mobile society, the provision of voice and data communications to a person away from his/her wireline telephone has become a major communications frontier. The early penetration of this frontier has been based on very limited portable communications approaches, e.g., cordless telephones, mobile radio telephones, and radio paging. Each of these approaches only partially satisfies portable communications needs. This paper describes an approach to providing universal digital portable communications integrated into telephone networks. A system configuration employing time-division multiple-access radio link architecture and frequency reuse is described. Issues affecting radio link transmission rates and radio system coverage are discussed. Characteristics and parameters of a possible system to supplement the wire (or fiber) loop are indicated.

I. Introduction

IN the highly mobile society of modern industrial nations, the provision of voice and data communications to a person away from his/her wireline telephone has become a major communications frontier [1]. One does not need to conduct a detailed market study to observe the large demand that exists for portable communications capability. The great popularity of cordless telephones, mobile radio telephones, radio paging, and other emerging portable communications technologies clearly demonstrates this intense pent-up demand. In the United States alone, about 20 million cordless telephones have been sold, and comparable sales have also occurred in many other countries. Sales of handheld portable sets for cellular mobile telephone systems are also increasing rapidly. The overall unrealized demand must be even greater than the proven demand for these emerging systems and features because they have only begun to penetrate this major frontier. That is, each of the current approaches to portable communications satisfies only a limited number of portable communications needs; none provides ubiquitous, universal, personal portable communications: This paper considers objectives, constraints, and a system approach for a universal digital portable communications system. The proposed system is not intended to replace vehicular systems.

The existing telephone network can be readily adapted to provide ubiquitous and universal portable communications to roving users. In order to do this, a network needs a widespread arrangement of fixed radio equipment connected to it as illustrated in Fig. 1. Adaptation is also needed to be able to identify people (i.e., portable sets) instead of identifying places (i.e., individual fixed telephones). This will require access to large databases for information about customer's services. It will also require rapid data transfer among switching equipment in different central offices. These network adaptations are beginning to occur now for other purposes; access to large databases is being used for providing special services. Also, a common channel signaling network (CCITT no. 7) is beginning to be implemented for rapid digital communications among central offices.

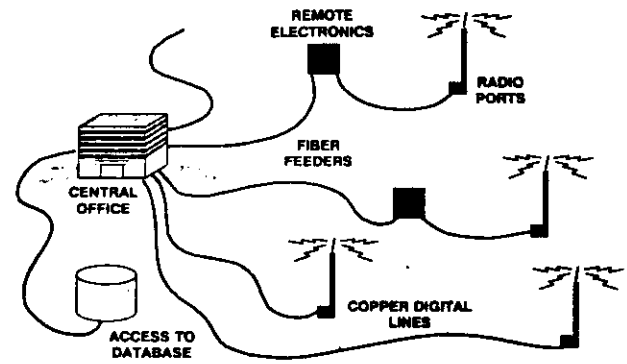


Fig. 1. Fixed radio equipment for providing universal digital portable communications as an integral part of a telephone network.

II. Limitations of Current Portable Communications Systems

It is useful at this point to consider briefly some examples of the limitations of existing approaches to portable communications. They are as follows:

A.

Current mobile systems (in North America: 150 and 450 MHz conventional and 850 MHz cellular) were designed to work with *vehicular* radio equipment. These designs have inherent characteristics that limit their suitability for portable handheld usage. Some of these characteristics are (see also [14] and [21]) as follows.

1) Vehicles are normally used *outdoors*. Portables are normally used where people spend most of their time—in *buildings*—where large-building penetration losses [5], [15], [16] add an average of about 20 dB (± a large stan-

Manuscript received August 25, 1986; revised February 2, 1987. This paper was presented at ICC '86, Toronto, Ont., Canada, June 1986.

The authors are with Bell Communications Research, Inc., Red Bank, NJ 07701.

IEEE Log Number 8714332.

0733-8716/87/0600-0764$01.00©1987 IEEE

dard deviation, which makes the 2σ level roughly 40–50 dB) relative to the median path loss encountered by an outdoor user. See the Appendix for further discussion of this point.

2) Since both a source of dc power (and its recharging) and a substantial stable mounting surface are available in a vehicle, its equipment typically can transmit an ERP of several watts. A portable unit with a lightweight, low-volume battery is not compatible with these power levels, and frequent recharging of its battery is not convenient. Current state of the art in battery design yields about a 30 dB natural difference in ERP between the two.

3) Body proximity and random orientation effects can reduce the portable's signal even further, relative to that of a vehicle.

4) Because vehicles are typically occupied only a fraction of the day, both average telephone traffic intensity (Erlangs per user in the busy hour) and total traffic per day for vehicular service are lower than that for personal portables, by up to a factor of ten.

5) Vehicles range over a wide area at high speed. It is natural, therefore, that the use of larger coverage areas (cells) is more important for vehicles in order to be economical; use of higher system ERP's and higher antenna masts produce the most cost-effective vehicular system design. However, this practice illuminates distant reflectors of RF energy, and delay spreads (echos) of up to tens of microseconds are often encountered in urban or hilly environments. This either precludes wide-band systems or makes adaptive delay equalization necessary.

6) The calling method used in current systems is far from ideal. A caller has to guess if and when the vehicle is occupied. The number of a vehicle, rather than that of a person, is used. The user must be in the vehicle (or near a compatible portable unit which is powered on) and must answer using it. Traffic intensity is unnecessarily increased by futile attempts.

7) The nature of vehicular movement and traffic density also lends itself to a different networking arrangement. Exchange areas for vehicles can be on the order of 1000 square km, while urban wireline exchanges (wire centers) are much smaller. While current vehicular systems have very few network access points, a dense personal portable system should be integrated into an exchange network as part of the distribution loops to the closest wire centers.

In summary, vehicular units and personal portables (handhelds) have many inherent differences in usage and in capability. Because of these differences, it is appropriate for vehicular and portable systems to be separate, yet to be coexistent.

B.

Cordless telephone handsets are designed to be radio extensions of wireline mainstations. They operate at low radio power levels and are battery efficient. However, they, too, have significant limitations.

1) A cordless handset is generally restricted to be used with its own base unit. Thus, a handset service area extends for, at most, 100–200 m around a particular base unit in a residential area; within large buildings, a 20–30 m range is more typical, depending on construction. The fixed association with one base unit location is a major disadvantage, limiting its utility, particularly to business users.

2) This approach does not save the investment in distribution facilities; the per-user wireline loop, including house/building wiring, is still required. In commercial locations, reassigning offices and work stations would still require shuffling telephone number assignments. Using cordless sets in a personal mode would require multiple loops to a residence and multiple base units.

3) Because of limited channel availability (ten channels in the United States) and because there is no coordination of channel usage, the density of handsets in an area is limited by the typical user's tolerance of crosstalk, mutual interference, and cutoff calls. Attempts to use cordless telephones in high-density housing or apartment buildings can result in unpredictable and out-of-control cochannel interference. Some new cordless telephones scan multiple channels (about 50) and measure the level of interference in each channel before selecting one to use for a call. This approach offers some improvement in cochannel interference, but high set density can result in unacceptably large numbers of calls interrupted as new users access in-use channels [20]. These multichannel cordless handsets and base units have most of the complexity (e.g., multichannel synthesis, signal quality measurement, sophisticated handset–base unit signaling, and channel assignment logic) needed to provide the more universal portable communication discussed later in this paper; but they still have the limitations noted above. That is, they are still cordless extensions, with only the simple phone cord replaced by a sophisticated radio link.

C.

Radio paging provides widespread inexpensive, lightweight, but limited communications in only one direction. Its limitations as a universal loop-equivalent system are obvious.

D.

Automatic Call Forwarding implemented in stored-program-controlled electronic switching systems provides communication away from a particular telephone, but this feature is still limited to a tethered connection and it requires continual user/system interaction.

III. Objectives for Universal Digital Portable Communications

Major goals of this system are:

- provision of the desirable and often necessary function of being completely tetherless and portable *anywhere*,
- replacement of the capital- and maintenance-intensive per-user end portion of the local distribution plant,

• utilization of the current Public Switched Telephone Network (PSTN) or planned Integrated Services Digital Network (ISDN) to the fullest extent possible.

From these follow a more detailed list of objectives:

1) efficient usage of scarce radio frequencies to serve a clear public need,

2) quality and reliability equivalent to wireline telephone and data service,

3) ubiquitous integrated services using tetherless portable telephones and data terminals,

4) privacy and security equivalent to wireline telephones,

5) small, convenient, and user-friendly portable sets with maximum time of use between battery recharging or replacement,

6) economical service at costs comparable to wireline telephone and data,

7) minimized radio frequency radiation hazard,

8) compatibility with the environment, i.e. minimized problems associated with the siting of fixed stations,

9) flexibility to accommodate different user needs and future technology advances, and

10) ability to accommodate a reasonable degree of user motion.

Objectives 5) and 7) necessitate low-power portable sets. Low power further implies short range and small coverage areas. Objective 8) implies low antenna masts, comparable to street lights and flag poles. These objectives are discussed in more detail in [21].

IV. System Configuration for Universal Digital Portable Communications

In a telephone network, telephone lines from central offices are grouped into cables or multiplexed onto single wire pairs that run via feeder cables to distribution points. (The cables may be replaced by glass fiber lightguides in the future.) From the distribution points, loop circuits proceed as separate wire pairs to telephone sets. For providing universal portable communications, the separate circuits at the ends of the telephone loops can be provided by using low-power demand-assigned digital radio links for the last 300–500 m. Radio could replace the following components of local loops:

- building wiring
- protector block
- drop wire
- per-user portions of the feeder and distribution plants
- splicing, both outside and at wire centers.

Such an implementation of loop radio links in residential areas is illustrated in Figs. 1 and 2. The fixed radio equipment must be configured to reuse radio frequencies efficiently and the equipment must be controlled by the network on a call-by-call basis to limit cochannel interference.

Proposals also are being made to distribute wide-band services such as video directly to businesses and residences using glass fibers. The demand for tetherless portable communications would not be affected by the availability of wide-band services. This is evident in that today almost everyone in the United States has a wireline telephone; but, nevertheless, 20 million cordless phones have been bought to connect to the ends of the copper loops to provide very limited tetherless communications. Fiber is not different from copper in being an unwanted tether. Therefore, wide-band fixed (tethered) services will complement, not compete with, tetherless portable communications services; freedom and wide bandwidth are not cross elastic.

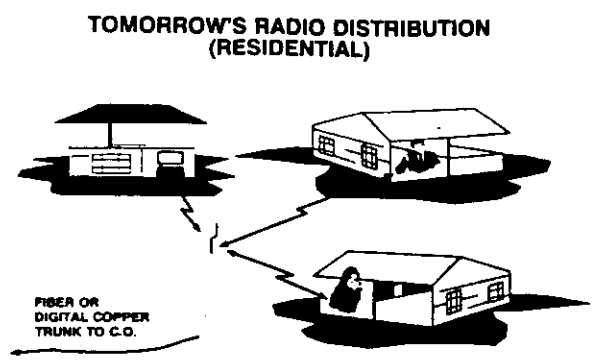


Fig. 2. A multiuser radio port implementing radio links for the ends of telephone distribution facilities.

A. Time-Division Multiple-Access Radio Link Architecture

Many portable handsets and data terminals could be served from one installation of fixed radio equipment at the end of a feeder line from a central office. The radio links would be shared among customers on the basis of individual user activity; that is, the radio links would be demand assigned. The fixed radio equipment would function as a radio port, i.e., an entry/exit point to the network via radio. Of course, a shared network radio port requires either multiple radio channels and multiple low-power radio transmitters and receivers or a time-multiplexed multiple-access multiple-channel technique. The latter approach can easily provide a flexible radio link configuration that is attractive for providing different user transmission rates.

In this approach, several fixed-rate bit streams would be time-division multiplexed (TDM) onto a radio carrier for transmission from a port to several active portable sets. (For a similar radio link architecture, see also [2].) Such a time-multiplexed plan is illustrated in Fig. 3(a) where the *H* indicates bits comprising a TDM frame header, and the numbered time slots indicate blocks of bits from fixed-rate bit streams. An average bit rate as low as 4 or 8 kbits/s could be a logical choice for the lowest multiplexed bit rate contained in a block. A portable voice or data set could then be assigned transmission rates in increments of the lowest rate by using more than one of the time-multiplexed blocks.

In the reverse direction, transmissions from portable sets to ports would be time sequenced and synchronized on a common frequency for time division multiple access

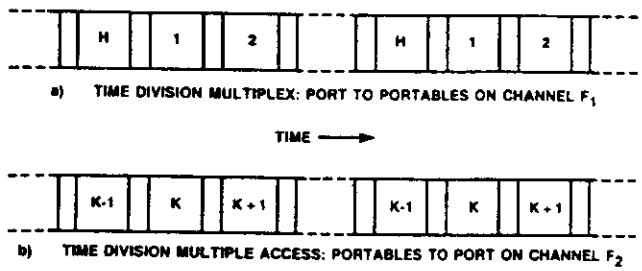


Fig. 3. Time-Division multiple-access radio link architecture.

(TDMA) as also indicated in Fig. 3(b). K represents an offset of an integer number of time slots, the offset eliminating the need for the portable to transmit and receive simultaneously (i.e., no need for a diplexer). Again, these transmissions would be quantized into some minimum average bit rate and a portable set could transmit at a multiple of the minimum rate by using more than one transmission time slot to the port. Thus, although the transmission rate on the radio links would be constant, a choice of user rates could be accommodated by using multiple blocks of transmission time.

In order to minimize the voice transmission delay and the change in the radio channel between TDMA frames, it is desirable to keep the frame period small. On the other hand, extra bits are needed in each frame for synchronization, signaling, and channel control, and a dead time is needed between blocks to prevent overlaps between transmissions from different portable sets caused by different propagation path delays or by TDMA synch misalignments. Therefore, to minimize the relative amount of this unproductive overhead, the frame period should be made large. A good compromise between these conflicting goals appears to be in the range of 10–20 ms; for an average bit rate on the order of 8 kbits/s, this results in about 100–150 information bits in a block.

A few different voice coding techniques operating at different average bit rates and providing differing inherent voice quality could be supported by the radio system, allowing a price/performance selection to be made by the customer. The addition of or change to a new speech coding technique supported by the fixed network could be accomplished by changing only software or circuit boards; the effects on the overall radio system would thus be minimal. Similarly, data customers could choose and pay for the digital data rates that they need. One or more of the multiplexed bit streams could be used to provide a contention channel for packet data. Transmission compatible with Integrated Service Digital Network (ISDN) rates could be supported using multiples of the minimum radio-link bit rate.

The maximum possible transmission rate for an individual portable radio link will be limited by the maximum spread in time delays that occurs for the multiple propagation paths resulting from reflections and scattering off walls and objects. An example of a "power versus delay" profile measured within a building is shown in Fig. 4.

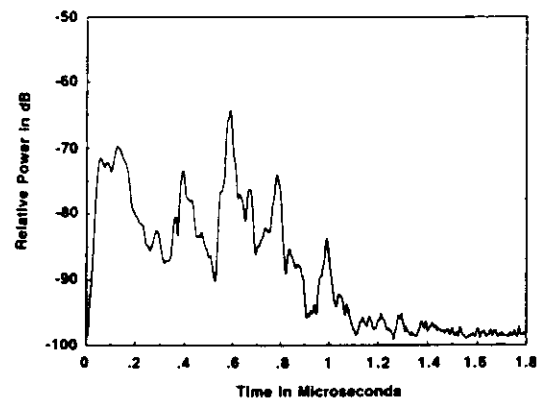


Fig. 4. A measured profile of average received power versus time delay at 850 MHz.

Such multipath propagation measurements [3] made in several office buildings, individual houses, and other portable communications environments suggest [4] that, for four-level digital phase modulation (QPSK), a maximum radio link bit rate in the range of 300–400 kbits/s could be transmitted without adaptive equalization in building environments. At locations that have an unusually large spread in transmission delay, antenna diversity, as discussed in Section IV-C, may be needed to maintain acceptably low error rates at the higher bit rates [4]. Allowing for overhead bits in headers and for encryption and system control, this would permit ISDN-compatible transmitted information rates of up to 144 or 288 kbits/s, by using many time slots. Radio channels could be about 125 or 250 kHz wide, with channel spacings of about 150 or 300 kHz.

TDM systems with low power and low antennas hold the promise of a more economical system design, relative to conventional methods, for the following reasons: 1) significantly fewer fixed station transceivers, 2) a simpler and faster channel synthesizer (at the portable), 3) simpler port antenna combining, 4) fewer intermodulation products, 5) no simultaneous transmit and receive at the portable unit (hence no need for a diplexer), 6) lowered probability of lightning damage, and 7) ease of maintenance.

Trunking efficiencies would be realized from the traffic concentration in the multiple-user port equipment. Therefore, for high market penetration levels, fewer user channels and less fixed radio equipment would be needed at a port than are needed in current cordless telephone installations that use a fixed base unit for each handset. Concentration factors of at least 3 or 4 are reasonable for ports serving 50–100 portable sets [19], depending on blocking objectives and user demand.

The user would need a somewhat more complex multichannel handset or data terminal than is needed for today's typical cordless telephone. However, the added set capability could provide the opportunity to free the user from the association with a wired connection and a specific base unit. The user then would have the opportunity to use the *same* portable set anywhere that fixed radio ports

are installed, i.e., at home, at the office, at a shopping center, or wherever a user wants to make or receive calls or data transmissions. The added set complexity is of the type that lends itself well to economical large-scale circuit integration. The small added cost for increased capability would tend to be offset by cost reductions that would result from eliminating the maintenance and installation of per-user wiring and from reduction in the amount of fixed radio equipment needed. A widespread low-power portable radio system would also provide a convenient and attractive means for calling-for-help for any problem, i.e., for fire or for threat of physical violence.

B. Privacy/Security

Voice privacy without degraded voice quality (and also security of user identification) can be readily provided by encrypting digital radio links. In contrast, voice privacy cannot be provided effectively on the analog radio links used for existing cordless telephones and mobile radio without compromising cost and quality goals. Rapid access to large databases through common channel signaling networks would provide a convenient mechanism for distributing encryption keys to different ports as needed for serving roving users.

C. Radio Coverage Issues

At network ports in residential areas, antenna heights could be about 7–10 m above ground level, at the height of streetlights and flagpoles. Estimates indicate that residential area ports separated by about 600 m could provide nearly continuous radio coverage for 5 mW portable handsets. Port antennas would be small and inconspicuous to minimize environmental concerns. Using a low port antenna also tends to limit the time delay spread.

The total frequency spectrum allocated to an arrangement of ports would be divided into sets of time-division multiplexed channels. The TDM channels or sets could be reused at ports sufficiently far apart for the cochannel interference to be at an acceptably low level. Such frequency reuse is indicated in Fig. 5 where different letters indicate different TDM channels or channel sets. Structured frequency reuse is necessary to achieve efficient radio spectrum utilization.

In order to provide universal ubiquitous service, radio coverage from network ports should also be available within large buildings [1], [21]. Coping with the severe excess attenuation [5], [15], [16] when trying to penetrate very large buildings would probably require moving the network radio port antennas inside the buildings; this in turn would permit reuse of frequencies within a building to serve the very large customer densities that occur in such environments.

Determination of network port radio coverage areas and frequency reuse intervals is complicated considerably by the vagaries of radio propagation within and around houses and buildings. The multipath propagation resulting from reflections and scattering from walls manifests itself as large variations in signal level over small-scale

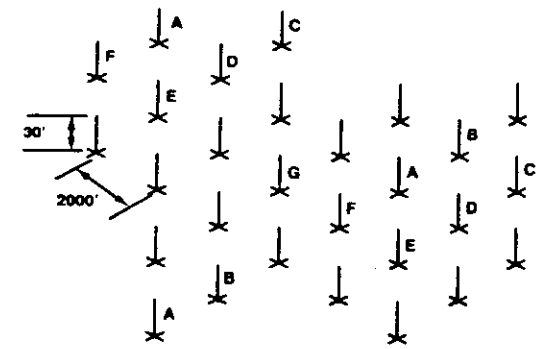


Fig. 5. A spatial arrangement of radio ports, illustrating frequency reuse (quadrangular grid).

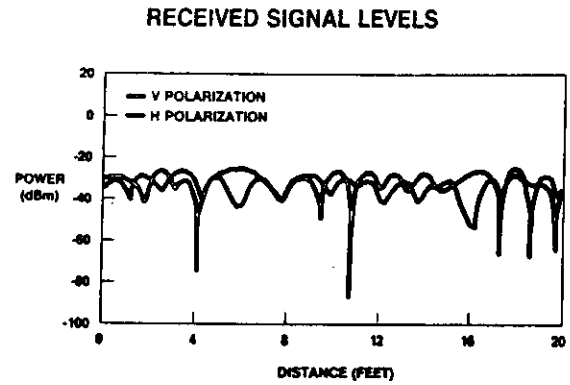


Fig. 6. Measured 800 MHz multipath signal variations.

distances on the order of a half wavelength. In the 1 GHz region, deep signal nulls can occur about every 10–20 cm because of random out-of-phase signal cancellation. Antenna diversity at either end of the radio link can reduce the small-scale multipath effects [7]. Antenna diversity takes advantage of the fact that two antennas separated by at least half of the average null spacing or having different polarizations [13] are not likely to be in signal nulls simultaneously. Fig. 6 illustrates simultaneously measured 850 MHz signal variation received within a building on vertically (V) and horizontally (H) polarized antennas from a moving vertically polarized transmitter [13]. The low correlation between the fluctuations on the two signals is obvious; the deep nulls do not occur simultaneously. Signal level distributions obtained in [7] indicate that the margin needed to allow for multipath propagation can be reduced to about 5 dB when antenna diversity mitigates the multipath variations.

Another signal impairment for portable radio links could result from random orientation of a handset. This could occur because of polarization misalignment of the randomly oriented handset antenna. Fortuitously, the random orientation effects are, however, mitigated by inherent coupling between polarizations in the multipath propagation medium [7], [8], [13].

Large-scale signal variations also occur from room to room, from building to building, and from house to house because radio paths are blocked or shadowed by walls,

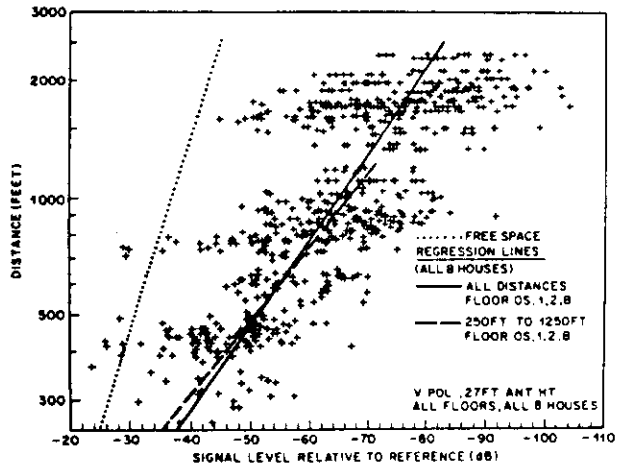


Fig. 7. Large-scale 800 MHz signal variations measured within and around eight houses.

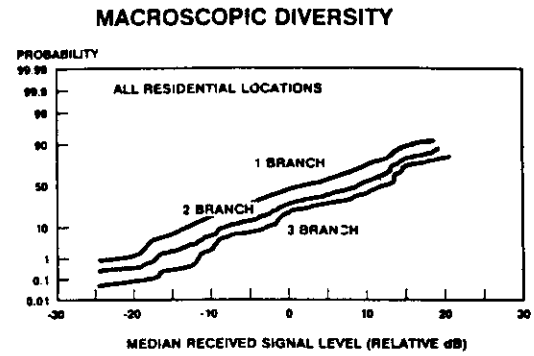


Fig. 8. Cumulative distributions of 800 MHz multipath signal medians, illustrating measured effectiveness of macroscopic selection diversity.

trees, terrain features, and other objects [6], [9]. Large-scale 850 MHz signal variations are illustrated in Fig. 7 where each data point represents the median signal level for small-scale multipath variations in residences, within a 1 m square area at about 1.5 m above the floor. A 9 m high port antenna was placed at various locations about the test area. The data were taken in different rooms on first and second floors, in basements, and immediately outside of eight houses. The solid regression line indicates a decrease in signal inversely as distance to the 4.5 power. The standard deviation σ of the log normal variation around the regression line is 10 dB, and at 450 m, the average excess attenuation relative to the theoretical free space value is 32 dB.

A portable set at maximum distance from a port in a residential area or in a large building will actually be at a comparable distance from several other ports (see Fig. 5). Random path blockage that causes low signal levels between a portable set and one port is not likely to exist between the portable set and the other ports. Thus, selection of a radio link to the port that yields the best signal quality would provide a form of large-scale or macroscopic diversity to mitigate large-scale signal variations [10], [11]. Fig. 8 shows measured cumulative distributions of median signals illustrating performance for two and three branch macroscopic selection diversity measured in and around eight houses [10]. A model based on these measurements indicates that 99 percent of users in the worst locations at maximum distance from four ports could be served if a 10 dB margin above the average were provided. Without macroscopic diversity, a margin of about 25 dB would be needed.

Diversity is a cost-effective system technique [17]. A block schematic of a possible combining concept is shown in Fig. 9 (for simplicity, only two branches are shown). The approach is conventional in physical layout, but could use error-detecting coding [12] for the diversity selection.

Portable-to-port radio link parameters for residential areas are summarized in Table I; the table is based on typical equipment parameters [1], [21] and the propagation margins discussed above. Because power is less constraining at the port transmitter, the reciprocal direction could be designed with a larger power margin.

The statistics of radio propagation variations also strongly affect the number of channel sets needed to limit cochannel interference to acceptable levels. Results from computer simulations [11] suggest that, with macroscopic diversity, the number of channel sets needed in residential areas is in the range of 25–36 sets, to provide a 16 dB signal-to-interference ratio over 99 percent of the coverage area. A 16 dB signal-to-interference ratio is the sum of the 5 dB multipath margin and the 11 dB signal-to-noise threshold from Table I. Although the actual necessary signal-to-interference threshold is not yet known, it will likely be comparable to the required signal-to-noise threshold. Again, without macroscopic diversity, the number of channel sets needed would exceed 100.

Propagation characteristics within large buildings vary widely from building to building, depending on construction materials and configurations [6], [18]. Because of the limited extent of most buildings, it is often not possible to obtain enough data to determine meaningful distance dependencies or other statistical parameters. From available data [6], [18], it appears that the average extra attenuation between floors can range from a few dB to 25 dB or more. Thus, in some buildings, frequencies could be reused every other floor, while in others, reuse might only be possible every five or six floors. Large differences also will exist for acceptable horizontal frequency reuse because of the similar large horizontal variability caused by metal in walls.

D. Call Processing and Network Issues

1) The intent is to make Universal Digital Portable Communications fully integrated with the PSTN/ISDN, while at the same time introducing a personal quality to the service. The calling signal must be some form of radio paging. A method under investigation has evolved from a service known as Directed Call Pickup, using CCITT No.

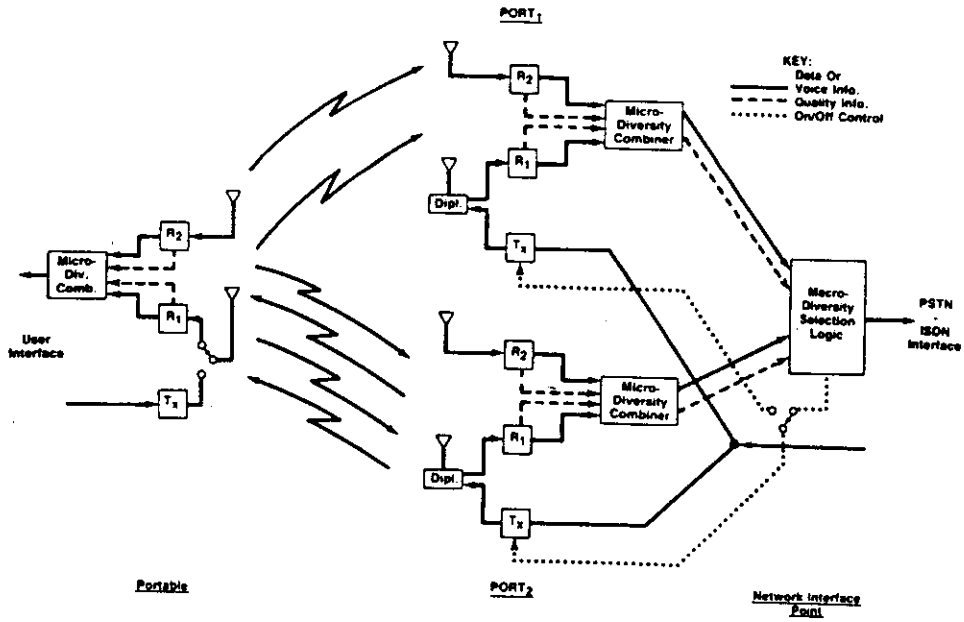


Fig. 9. Block diagram of diversity combining.

TABLE I
SUMMARY OF RADIO LINK PARAMETERS (PORTABLE TO PORT)

kT in 1 Hz BW	−174 dBm	
Noise figure (Port receiver)	+3 dB	
Noise BW^a (16kb/s, 4 level)	+40 dB	
S/N threshold (10^{-3} error rate, coherent)	+11 dB	
Port receiver sensitivity	−120 dBm	
Port antenna gain		+8 dB
Portable antenna gain		−4 dB
Portable transmit power (average)[a]		+7 dBm
Total radio link gain		131 dB
Free space attenuation (450 m, 900 MHz)		84 dB
Large-scale average excess attenuation (450 m)	32 dB	
Multipath margin with antenna diversity	5 dB	
Margin for large-scale variation (4 br. macro diversity)	10 dB	
Margin for excess attenuation (450 m)	47 dB	47 dB
Total propagation attenuation		131 dB

[a]Note: Peak power (TDMA transmission) is in the range of 20 dBm. Transmit power and noise bandwidth are normalized to average power and average transmission rate, rather than the peak TDMA values.

7 common channel signaling. The user would be paged via either a) one of several standard paging systems or b) a special paging system designed for and colocated with the ports. If the user chooses to answer using a portable personal telephone, the set would be energized, identification would be automatic, and CCITT No. 7 would transmit information between the originating and the answering central office to set up the connection. The Call Forwarding feature is an inherent part of the design.

2) To minimize channel holding time, signaling via Pre-Origination Dialing will be employed; that is, the user of the handheld unit will enter the number being called into the unit's memory prior to system seizure.

3) Supervision will be continuous during a call. In this

way, a set which malfunctions or is taken into a high loss area during a call does not leave the connection held up unnecessarily.

4) For several reasons, the reliability of equipment in such a system can be expected to be high.

- Low-power convection-cooled circuits are inherently more reliable (longer mean time to failure) than their high-power-dissipation counterparts.
- Low antennas are less vulnerable to lightning and to wind.
- Installations which are easily accessible (i.e., low) have shorter mean time to repair.
- Small, uncomplicated ports could easily be replaced *in toto*, then repaired at a central facility, especially if channel assignment and other site-specific parameters can be remotely controlled by system software.
- Complete failure of one port or one entire diversity group partially degrades but does not completely eliminate service in an area.

E. Typical Radio System Characteristics

In summary, parameters and characteristics of a possible system are as follows.

- Probability of coverage within service area—greater than 0.99.
- Portable transmitter average power—5 mW.
- TDM/TDMA with QPSK modulation.
- Structured frequency reuse; demand assigned channels.
- Diversity against both small- and large-scale signal variations.
- 16 kbit/s speech coding rate; voice delay less than 30–50 ms.
- Flexible system bit rates in 8 kbit/s increments.
- Channel information rate—in the range of 144–288 kbits/s.
- Channel spacing—between 150 and 300 kHz.
- Operating frequency band—somewhere between 400 and 4000 MHz.
- Built-in privacy.
- Residential port separation—up to 600 m.
- Residential port antenna height—7–10 m.
- In-building port separation—up to 60 m.

V. Future Directions

Digital radio combined with extensive network control has not been used before for portable communications in such a difficult radio environment. Therefore, continuing research must determine the limitations imposed by the harsh radio environment around and within buildings, and must adapt or invent radio link and system control techniques that work, are simple to use, and are economical enough for universal acceptance. Also, a suitable portion of the radio spectrum must be allocated. While many questions must be resolved and equipment constructed before a feasible system can be demonstrated, initial research and study of system implications are quite promising.

Appendix
The Mixing of In-Building and Outside Radio Populations

Let us compare the median carrier-to-noise[1] ratio at an elevated base station receiver of a radiocommunications system for two cases:

1) from an on-street source: $\overline{C/N}]_{os}$
2) from an in-building source: $\overline{C/N}]_{ib}$.

In the past, it has been found useful to define a building penetration loss L_{bp}: the ratio of in-building received carrier power to on-street received power in the same city block from the same base station transmitter. This has been measured by several experimenters [5], [9], [15], [16] to have a log-normal distribution, with a median value near ground level of about 20 dB and with a wide variance (σ_{bp} is about 8 dB). L_{bp} is clearly affected by the floor of the building, the construction, the surroundings, etc. Its median value decreases with height.

Reciprocity holds so that on a given transmission path and a given frequency, the base transmitter-to-remote portable receiver path loss equals the portable transmitter-to-base receiver path loss. Let us also assume that in-building and on-street sources have the same ERP. It then follows that $\overline{C/N}]_{ib}$ is less than $\overline{C/N}]_{os}$ by L_{bp}.

To illustrate the effect which L_{bp} has on the cumulative probability distribution (CPD) of signal strength, refer to Fig. 10. There we show two log-normally distributed CPD's; the one labeled *OS* applies to on-street users, and the one labeled *IB* applies to in-building users. *OS* has a mean μ_{os} arbitrarily chosen to be 0 dB and a standard deviation σ_{os} of 8 dB, while *IB* has μ_{ib} equal to -20 dB and σ_{ib} of 10 dB (losses are independent, such that $\sigma_{ib}^2 = \sigma_{os}^2 + \sigma_{bp}^2$; thus, σ_{bp} has a conservative value of 6 dB, for this example). From these, it can be seen that, if indoor users are served on a system originally designed to give good service to 90 percent of the outside users at some chosen radius, then only 16 percent of the in-building users will receive that same service. (In practice, many of the call attempts from in-building units will fail, so that the probability of good service (given that it was possible to start the call at all) will be higher.)

On the other hand, if a 50/50 mix of indoor and outdoor users is taken as the design population, the CPD labeled 50/50 MIX must be used. In this case, the following hold.

- The CPD is no longer log-normal, but takes on a bimodal character.

[1]The term "noise" is used here to include random Gaussian noise, cochannel interference, impulse noise (man-made or natural), adjacent-channel interference, etc. Note also that the Rayleigh-distributed multipath fading, which occurs on the scale of fractions of a wavelength, is averaged out for this comparison.

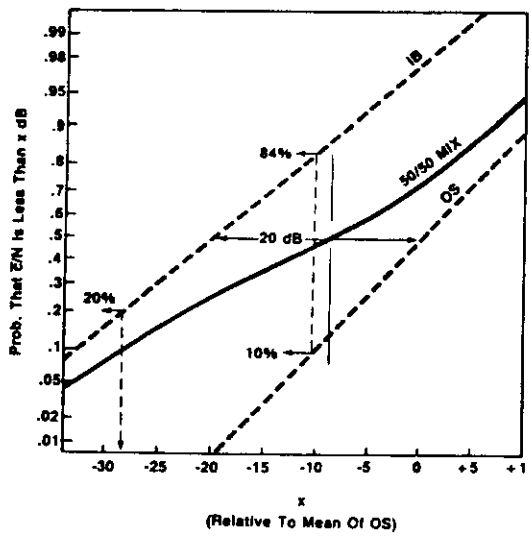


Fig. 10. Cumulative distributions of $\overline{C}/N$ for in-building, on-street, and 50/50 mixed populations.

- The spread between the 10th and 90th percentiles has increased by about 10 dB.
- The population below the median will be overwhelmingly made up of inside units. And if the 10th percentile is the design criterion for the system, roughly 99.98 percent of the outside units will exceed the design objective; thus, the system is overdesigned for the outside units.

It is important to remember that this comparison assumed that the indoor units were capable of the full ERP of the outdoor units and differed only in their environment. When the practical reality of a lower ERP in handheld units is taken into account, the comparison becomes even more skewed.

Decreasing the effective radiated power (ERP) of outside units below that of the inside portables via some sort of automatic power control algorithm might help matters somewhat, provided that the algorithm is sufficiently accurate and very fast acting, yet does not respond to the multipath fading. So far, this has yet to be demonstrated in an interference-impaired situation.

References

[1] D. C. Cox, "Universal portable radio communications," *IEEE Trans. Veh. Technol.*, pp. 117–121, Aug. 1985.
[2] K. Kinoshita, M. Hata, and K. Hirade, "Digital mobile telephone system using TD/FDMA scheme," *IEEE Trans. Veh. Technol.*, vol. VT-31, p. 153, Nov. 1982.
[3] D. M. J. Devasirvatham, "Multipath time delay jitter measured at 850 MHz in the portable radio environment," *IEEE J. Select. Areas Commun.*, this issue, pp. 000–000.
[4] C.-I. Chuang, "The effects of time delay spread on portable radio communications channels with digital modulation," *IEEE J. Select. Areas Commun.*, this issue, pp. 000–000.
[5] E. H. Walker, "Penetration of radio signals into buildings in the cellular radio environment," *Bell Syst. Tech. J.*, pp. 2719–2734, Nov. 1983.
[6] S. E. Alexander, "Characterizing buildings for propagation at 900 MHz," *Electron. Lett.*, p. 860, Sept. 28, 1983.
[7] D. C. Cox, "Antenna diversity performance in mitigating the effects of portable radiotelephone orientation and multipath propagation," *IEEE Trans. Commun.*, vol. COM-31, pp. 620–628, May 1983.
[8] D. C. Cox *et al.*, "Crosspolarization coupling measured for 800 MHz radio transmission in and around houses and large buildings," *IEEE Trans. Antennas Propagat.*, pp. 83–87, Jan. 1986.
[9] D. C. Cox, R. R. Murray, and A. W. Norris, "800 MHz attenuation measured in and around suburban houses," *Bell. Labs. Tech. J.*, July–Aug. 1984.
[10] H. W. Arnold, R. R. Murray, and D. C. Cox, "Macroscopic diversity in the portable radiocommunication environment," presented at the 1985 North Amer. Radio Sci. Meet., Vancouver, B.C., Canada, June 17–21, 1985.
[11] R. C. Bernhardt, "Macroscopic diversity in frequency reuse radio systems," *IEEE J. Select. Areas Commun.*, this issue, pp. 000–000.
[12] L.-F. Chang, H. W. Arnold, R. C. Bernhardt, and P. T. Porter, "Coding as a means to implement diversity selection in a frequency-reusing portable radio system," in *Conf. Rec. 2nd Nordic Seminar Digital Land Mobile Radio Commun. (DMR II)*, Stockholm, Sweden, Oct. 14–16, 1986, pp. 90–93.
[13] S. A. Bergmann and H. W. Arnold, "Polarization diversity in the portable communications environment," *Electron. Lett.*, pp. 609–610, May 22, 1986.
[14] J. Engle and S. Thro, "Frequency reuse boosts RF/data terminal throughput," *Mobile Radio Technol.*, Feb. 1985.
[15] L. P. Rice, "Radio transmission into buildings at 35 and 150 mc," *Bell Syst. Tech. J.*, pp 197–210, Jan. 1959.
[16] J. M. Durante, "Building penetration loss at 900 MHz," in *Conf. Rec., 1973 IEEE Vehic. Technol. Conf.*, pp. 1–7.
[17] P. T. Porter and R. C. Bernhardt, "Cost comparisons of macroscopic diversity architectures for universal digital portable communications," in *Proc. IEEE GLOBECOM*, New Orleans, LA, Dec. 1–4, 1985, pp. 1008–1012.
[18] R. R. Murray, H. W. Arnold, and D. C. Cox, "815 MHz radio attenuation measured within a commercial building," in *Proc. IEEE Antennas Propagat. Symp.*, Philadelphia, PA, June 1986, pp. 209–212.
[19] M. T. Manfred, G. A. Nelson, and C. H. Sharpless, "Loop plant electronics: Digital loop carrier systems," *Bell Syst. Tech. J.*, vol. 57, pp. 1129–1156, Apr. 1978.
[20] A. J. Motley and F. A. Al-Salihi, "Simulation of advanced cordless telephones," presented at the 3rd Int. Conf. Land Mobile Radio, Cambridge, England, Dec. 10–13, 1985.
[21] D. C. Cox, "Universal digital portable radio communications," *Proc. IEEE*, April 1987.

Donald C. Cox (S'58–M'61–SM'72–F'79), for a photograph and biography, see this issue, p. 762.

Hamilton W. Arnold (S'67–M'67–SM'85) received the A.B. degree in physics from Occidental College, Los Angeles, CA, in 1965, and the M.S. and Sc.D. degrees in electrical engineering from Columbia University, New York, NY, in 1967 and 1971, respectively.

From 1971 to 1983 he was employed by Bell Laboratories, Holmdel, NJ. His research activities there included millimeter-wave mobile radio and satellite propagation, advanced satellite communications systems, and digital transmission techniques in the frequency-selective mobile radio environment. In 1983 he joined Bell Communications Research Inc., Red Bank, NJ, where he manages a group conducting research on radio propagation and techniques.

Dr. Arnold is a member of Commission F of USNC/URSI and Sigma Xi.

Philip T. Porter (M'59-SM'84) received the B.A. and M.A. degrees in physics from Vanderbilt University, Nashville, TN, in 1952 and 1953, respectively.

He is a District Manager with Bell Communications Research, Inc., Red Bank, NJ. His group studies intrasystem frequency reuse and other efficiency improving techniques for future telecommunications systems meant to provide untethered personal communication in suburban and urban, in-building environments. Prior to 1984 he was employed by AT&T Bell Laboratories, where he was involved (1963-1970) in development of Bellboy paging, IMTS, the Metroliner system, and others (including early planning studies leading to 850 MHz cellular systems). From 1971 to 1977 he supervised a group involved in systems planning of network and mobile control logic for the Advanced Mobile Phone Service; later, he was responsible for long-range mobile radio studies and for international liaison.

Mr. Porter is a member of Phi Beta Kappa and of various IEEE Societies. He is also a U.S. delegate to CCIR Study Group 8A.

A medium with advantages over radio...and some drawbacks

Non-Directed Infrared Links for High-Capacity Wireless LANs

JOSEPH M. KAHN, JOHN R. BARRY, MALIK D. AUDEH, JEFFREY B. CARRUTHERS, WILLIAM J. KRAUSE, AND GENE W. MARSH

The emergence of portable information terminals in future work and living environments is expected to accelerate the introduction of wireless LANs. Such portable terminals should have access to all of the services that will be available on wired networks. Unlike their wired counterparts, portable devices are subject to severe limitations on power consumption, size, and weight. The desire for inexpensive, high-speed links satisfying these requirements has motivated recent interest in infrared wireless communication [1-5].

As a medium for short-range, indoor communication, infrared offers several significant advantages over radio, including a virtually unlimited spectral region that is unregulated worldwide. Near-infrared and visible light are close together in wavelength, and they exhibit qualitatively similar behavior. Both are absorbed by dark objects, diffusely reflected by light-colored objects, and directionally reflected from shiny surfaces. Both types of light penetrate through glass, but not walls or other opaque barriers. As a result, infrared communications can readily be secured against eavesdropping. Moreover, it is possible to operate at least one infrared link in every room of a building without interference, so that the potential capacity of an infrared-based network is extremely high. When an infrared link employs intensity modulation with direct detection (IM/DD), the short carrier wavelength and large, square-law detector lead to efficient spatial diversity that prevents multipath fading. By contrast, radio links are typically subject to large fluctuations in received signal magnitude and phase.

The infrared medium is not without drawbacks, however. In many indoor environments there exists an intense infrared ambient, arising from sunlight, incandescent lighting, and fluorescent lighting, which induces noise in an infrared receiver. In virtually all short-range, indoor applications, IM/DD is the only practical transmission technique. The signal-to-noise ratio of a DD receiver is proportional to the square of the received optical power, implying that IM/DD links can tolerate only a comparatively limited path loss. Often, infrared links must employ relatively high transmit power levels and operate over a relatively limited range. While the transmitter power level can usually be increased without fear of interfering with other users, transmitter power may be limited by concerns of power consumption and eye safety, particularly in portable transmitters. Some of the characteristics of infrared and radio indoor wireless links are compared in Table 1.

Using directional infrared transmitters and receivers, it is possible to achieve high bit rates and long link ranges using relatively modest transmitter power [6]. In most applications of wireless LANs, however, it is desirable to form links using omnidirectional transmitters and receivers, alleviating the need for careful alignment between them. This article will focus on such non-directed links. As illustrated in Fig. 1, non-directed infrared links may be classified into two categories: line-of-sight (LOS) and diffuse. LOS links depend upon the existence of an unobstructed path between transmitter and receiver. Diffuse links alleviate the need for a direct LOS path by relying on light scattered from a large diffuse reflector, such as a ceiling. Because it is difficult to block all of the light reflected from such a large surface, diffuse links are more robust than LOS links, and may be preferable for many applications.

Fig. 2 illustrates two different paradigms for creating wireless infrared LANs serving portable information terminals. When two or more portables are located in the same room, they may communicate directly with each other on a peer-to-peer basis, forming an ad hoc network. Portable transceivers designed for such ad hoc interconnection should consume little power and be relatively inexpensive. Alternatively, infrared links may also be used to connect portables to base stations that are interconnected by a wired backbone network. Such an installed network would permit portables to communicate with multimedia and compute servers, or with portables located in other rooms. In this scenario, the portable terminals should be inexpensive and low-power, but it might be permissible for the base stations to be more complex and to consume greater power. In some future high-performance multimedia wireless computing environments, the portable terminals may serve mainly as a human interface, accepting pen and keyboard input, but displaying full-motion video. The very high-capacity downlinks (tens of Mb/s per base station) and moderate-capacity uplinks (several Mb/s per base station) required of such a system would be particularly well-matched to the capabilities of infrared communication. Smaller rooms could be served by a single base station, while rooms larger than about 10 m x 10 m may require more than one base station. Techniques for accommodation of multiple base stations in one room will be touched upon below.

Despite a relative scarcity of research publications on wireless infrared communications, the technology has found wide commercial application. Directed infrared beams are commonly used in remote-control devices, as well as in serial links for

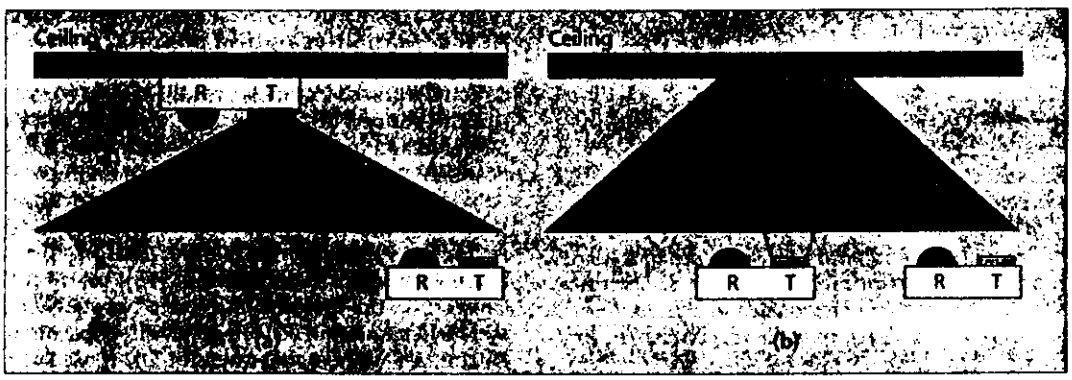
1070-9916/94/$04.00 © 1994 IEEE

computer peripherals operating at bit rates up to 112 kb/s [7]. Diffuse infrared has been employed for several years in commercial audio transmission systems. During the past year, there have been several new products using diffuse infrared transmission to permit interconnection of portable computers. Both IBM and Photonics are marketing modems[8,9] that permit ad hoc, peer-to-peer interconnection of notebook computers at a bit rate of 1 Mb/s, achieving coverage of a 10 m x 10 m room. These modems use light-emitting diode (LED) sources with 16-pulse-position modulation (PPM), and support a CSMA/CA protocol. Spectrix is offering portable terminals [10] that use 2-PPM of LED sources to achieve a 4-Mb/s transmission speed over a 15-m range. These links permit wireless communication with base stations connected to a backbone network, making use of a deterministic reservation/polling SNMP protocol. While there are currently no standard transmission formats or protocols for wireless infrared networks, a subgroup of the IEEE 802.11 committee is expected to draft the first standards for wireless infrared networks in early 1994 [11].

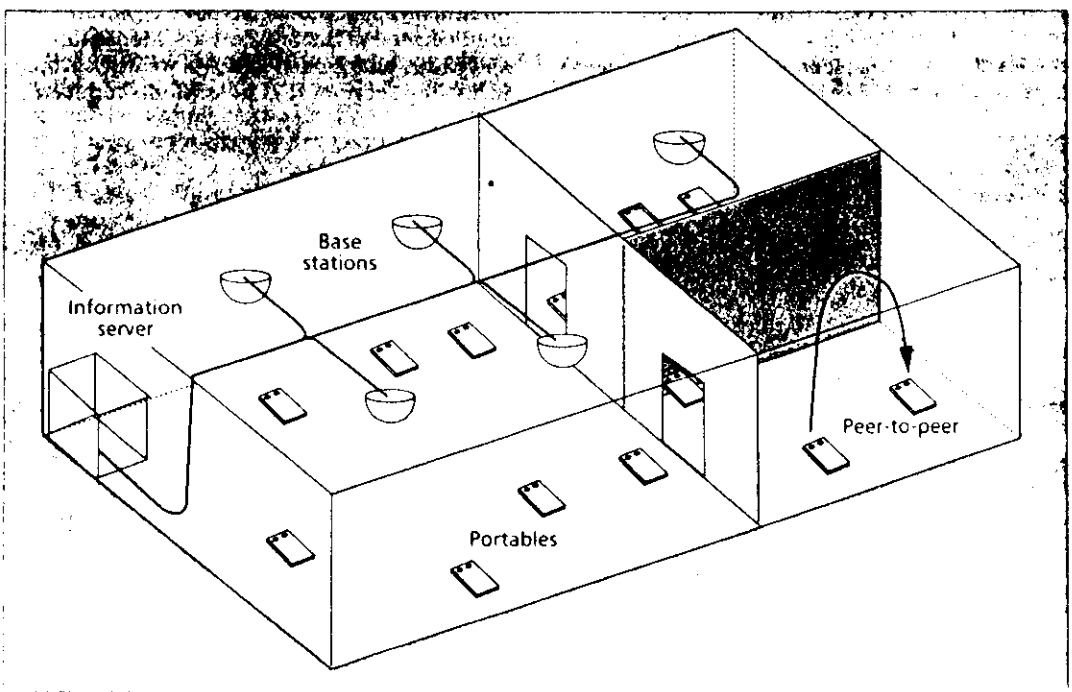
Channel Property	Non-Directed Infrared	Radio
Path loss	High	High
Multipath fading	No	Yes
Multipath distortion	Yes	Yes
Dominant noise	Shot noise from background light	Interference from other users
Input $X(t)$ represents	Power	Amplitude
SNR proportional to	$\|X(t)\|^2$	$\|X(t)\|^2$
Average power proportional to	$\int X(t)dt$	$\int \|X(t)\|^2 dt$
Bandwidth limitation	Photodiode capacitance and transit time	Regulatory

■ **Table 1.** *Comparison of infrared and radio for indoor communications.*

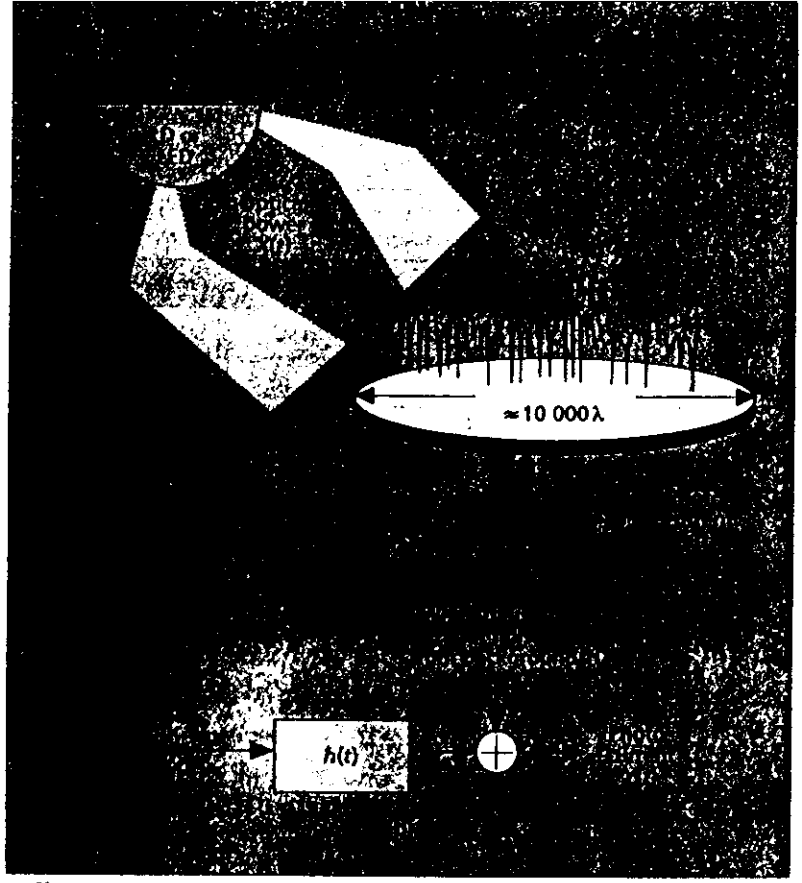
Using technology employed in current commercial systems, it should be relatively straightforward to extend the bit rate of non-directed infrared communication links to approximately



■ **Figure 1.** *Configurations of non-directed indoor infrared links: (a) line of sight, (b) diffuse. T and R denote transmitter and receiver, respectively.*



■ **Figure 2.** *Network of portable multimedia terminals using wireless infrared access to wired backbone. A peer-to-peer interconnection of two portables is also shown (from [22]).*



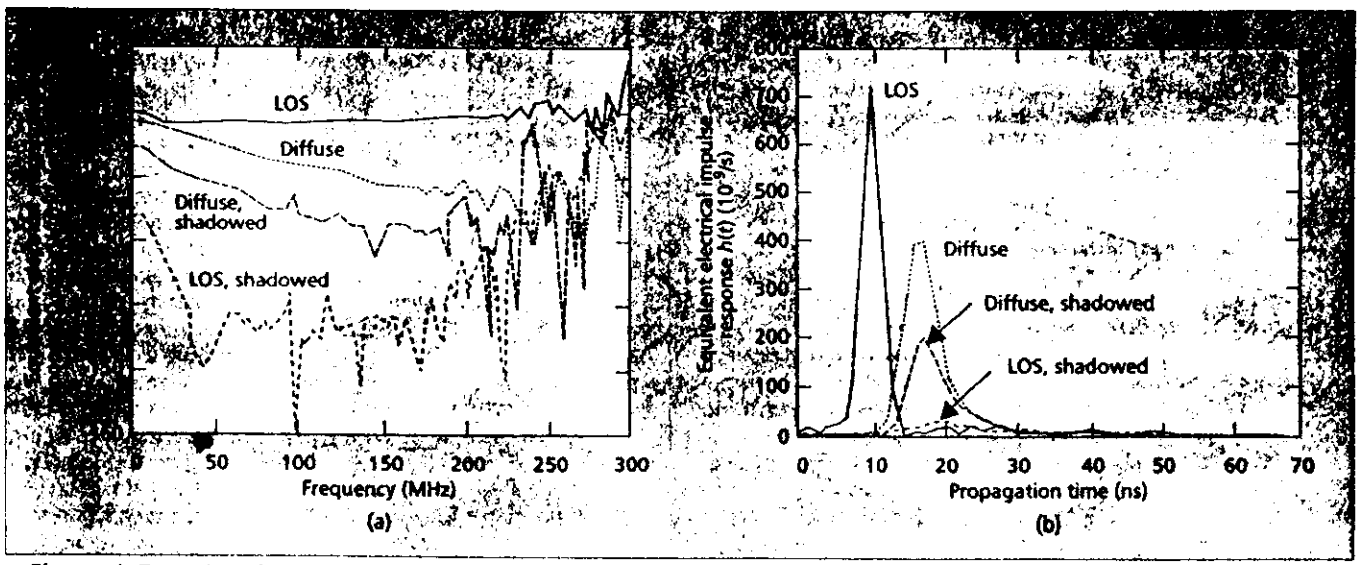
■ **Figure 3.** *Modeling non-directed infrared channels with intensity modulation and direct detection. (a) The received optical electric field generally exhibits spatial variation of magnitude and phase on the scale of a wavelength. The photocurrent is proportional to the integral over the detector surface of the optical power. (b) The channel can be modeled as a fixed, linear, baseband system with input X(t), output Y(t), impulse response h(t), and additive white, Gaussian noise.*

10 Mb/s. Higher transmission speeds will be desirable in future wireless computing environments. In this article, we describe the physical obstacles to achieving higher bit rates, and we discuss the technical means to achieve bit rates as high as 100 Mb/s.

Nature of Non-Directed Optical Channels

After non-directed propagation, an infrared beam typically consists of an unknown superposition of modes, making it very difficult to implement efficiently heterodyne or homodyne optical detection. Such techniques are also too complex for this cost-sensitive application, so that IM/DD is the only practical alternative. The modeling of non-directed infrared channels with IM/DD is illustrated in Fig. 3. The transmitted waveform $X(t)$ is the instantaneous optical power of the infrared emitter. The received waveform $Y(t)$ is the instantaneous current in the receiving photodetector, which is proportional to the integral over the photodetector surface of the total instantaneous optical power at each location.[1] As shown in Fig. 3(a), the received electric field generally displays spatial variation of magnitude and phase,[2] so that "multipath fading" would be experienced if the detector were smaller than a wavelength. Fortunately, typical detector dimensions are thousands of wavelengths, leading to efficient spatial diversity that prevents multipath fading. As the transmitted optical power $X(t)$ travels along various paths of different lengths, non-directed infrared channels are still subject to multipath-induced distortion. The channel can be modeled as a baseband linear system, with input power $X(t)$, output current $Y(t)$, and an impulse response $h(t)$, which is fixed for a given arrangement of transmitter, receiver and intervening reflectors. A mathematical derivation of this channel model can be found in [12], where it is shown that the linear relationship between $X(t)$ and $Y(t)$ is a consequence of incoherent propagation. By contrast, we note that when IM/DD is used with narrow-linewidth sources in dispersive single-mode optical fibers, the relationship between $X(t)$ and $Y(t)$ is nonlinear [13].

In many applications, non-directed infrared links are operated in the presence of intense infrared and visible background light. As we will see in the following section, it is possible to reduce this



■ **Figure 4.** *Equivalent electrical magnitude response (a) and impulse response (b) of non-directed infrared channels. Measurements were performed in an empty 5.5 m x 7.5 conference room having a ceiling height of 3.5 m (from [12]).*

received ambient light by optical bandpass or longpass filtering, but this background still adds a white, nearly Gaussian shot noise $n(t)$ that is the limiting factor in the signal-to-noise ratio (SNR) of a well-designed receiver. Our channel model is summarized by

$$Y(t) = X(t) \otimes h(t) + n(t), \quad (1)$$

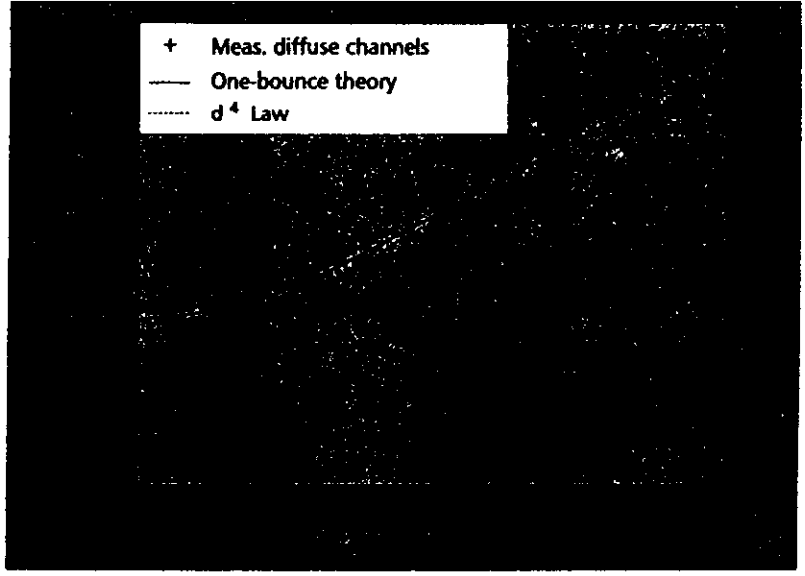
where the ⊗ symbol denotes convolution. As in linear electrical or radio channels with additive noise, the SNR is proportional to $|X(t)|$.[2] However, infrared channels differ from conventional channels because the channel input $X(t)$ represents power. Thus, $X(t)$ cannot be negative, and the average power is proportional to a time integral of $X(t)$,

$$P_{avg} = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} X(t)\, dt, \quad (2)$$

rather than the usual $|X(t)|^2$ which is appropriate when $X(t)$ represents amplitude.

While non-directed propagation alleviates the need for physical alignment between transmitter and receiver, a major drawback of this approach is the signal distortion caused by reflections from ceilings, walls, and other objects. To evaluate the possible impact of multipath distortion on high-speed infrared links, we have performed experimental characterization of about 90 channels in five offices and conference rooms [12]. We used a vector network analyzer to perform swept-modulation-frequency characterization [14] of the channel frequency response $H(f)$. The 832-nm transmitter emitted a Lambertian radiation pattern, i.e., with power-per-unit solid angle proportional to the cosine of the angle with respect to transmitter surface normal. Our receiver exhibited a 3-dB cutoff frequency of 150 MHz. During all measurements, the receiver was placed at desk height and pointed upwards. To form LOS configurations, the transmitter was placed near the ceiling and pointed straight down, while for diffuse configurations, it was placed at desk height and pointed straight up. Shadowed LOS and diffuse configurations were formed by having a person stand next to the receiver so as to block the dominant reception path. The channel frequency response $H(f)$ is scaled so that $H(0)$ represents the ratio of the optical power received by a 1-cm^2 detector to the total transmitted optical power. We obtain the channel impulse response $h(t)$ by inverse Fourier transformation of $H(f)$.

Fig. 4 presents the magnitude and impulse responses of four different non-directed infrared link configurations, measured in an empty conference room. While the details of channel responses depend on the link geometry, responses measured at all positions in all rooms exhibit qualitative similarity to Fig. 4. Unshadowed LOS impulse responses are dominated by a short initial pulse, and the strongest distinct reflections typically arrive 15 to 20 ns after the initial pulse. Dominance of the short initial pulse leads to magnitude responses that are flat at high frequencies. Unshadowed diffuse impulse responses exhibit a significantly wider initial pulse, which has a width of about 12 ns at 10 percent height, corresponding to the existence of a continuum of different path lengths between illuminated portions of the ceiling and the receiver.



■ **Figure 5.** *Optical path loss vs. horizontal separation between transmitter and receiver for diffuse, unshadowed channels. Measurements were performed in a 8.5 m x 9 m office having a ceiling height of 2.4 m, with transmitter and receiver arranged as in Fig. 1(b). Solid line is result of a model taking account only of the first reflection from the ceiling, which has a diffuse reflectivity of 80 percent. At large separation, results of that model approach a d^4 law (from [12]).*

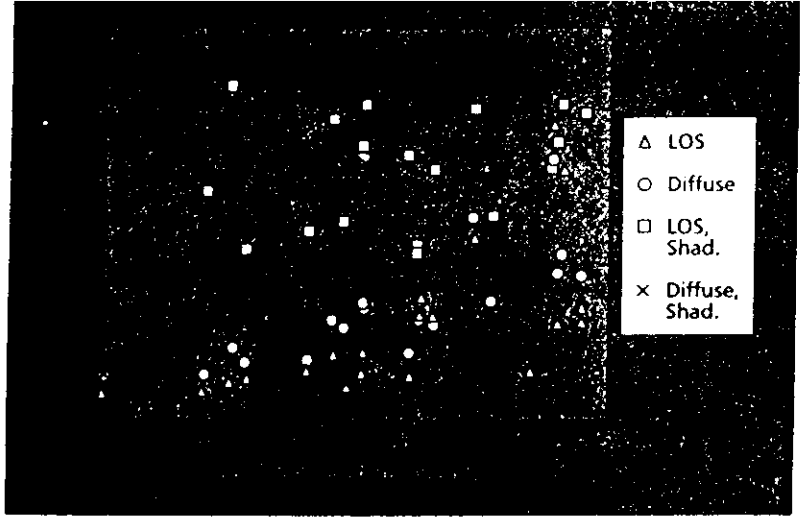
This continuous distribution of path delays leads to a steady decrease in the channel magnitude response at high frequencies. For all channels, the impulse response may contain significant energy as long as 70 ns after its initial non-zero excursion. The d.c. gain of all channels is enhanced over that at high frequencies, because it includes the contribution due to the entire duration of the impulse response.

Shadowed channels exhibit characteristics that are slightly less predictable than the unshadowed channels. The shadowed LOS impulse response typically resembles the corresponding unshadowed response with the dominant initial pulse removed, since only indirect propagation paths remain. We observe that in LOS configurations, shadowing significantly degrades the channel frequency and impulse responses. Diffuse configurations are far less vulnerable to shadowing than their LOS counterparts, because in diffuse configurations there exist many possible propagation paths between the illuminated ceiling area and the receiver. In diffuse configurations, shadowing produces a slight broadening of the impulse response, and a slight increase in the rate of falloff of the magnitude response with increasing frequency.

As the frequency response $H(f)$ is relatively flat at low frequencies, the equivalent d.c. electrical gain $H(0)$ is probably the single most important quantity characterizing a non-directed infrared channel. The same information is conveyed by the optical path loss, which we define as the ratio of the average transmitted power to the average power received by a detector of 1-cm^2 area, i.e., the reciprocal of $H(0)$. The path losses of diffuse channels measured in one office are shown in Fig. 5. These measured path losses can be fitted with reasonable accuracy by a model [15] that considers only the light undergoing one diffuse reflection from the ceiling en route from source to receiver. We note that for large horizontal separation between trans-

[1] *The detector is equivalent to a two-dimensional array of many antennas whose receptions are squared, lowpass filtered, and summed.*

[2] *Under very unusual circumstances — with point or plane-wave source, LOS propagation, and carefully aligned detector — this spatial variation would exhibit a regular pattern, but in typical cases of LOS or diffuse propagation, it appears to be random.*



■ **Figure 6.** *Root-mean-square delay spread versus horizontal separation between transmitter and receiver, measured in five rooms (from [12]).*

mitter and receiver, this model predicts that diffuse path losses should increase as the fourth power of horizontal separation. Considering measurements performed in five rooms, for a given horizontal separation, the unshadowed LOS configuration generally yields the lowest path loss, and the path loss of the corresponding unshadowed diffuse configuration is typically 1 to 3 dB higher. In the presence of shadowing, however, LOS configurations exhibit typically a 7 to 10 dB increase in path loss, while in the diffuse case, the corresponding increase is about 2 to 5 dB. As a result, shadowed diffuse configurations typically yields path losses that are 2 to 5 dB smaller than the corresponding shadowed LOS configurations, indicating the robustness of an extended, diffuse source.

In comparing the multipath dispersion of different channels, we will consider the channel r.m.s. delay spread. The delay spread is calculated from the impulse response according to:

$$r.m.s.\ delay\ spread\ of\ h(t) = \left[\frac{\int (t-\mu)^2 h^2(t)\,dt}{\int h^2(t)\,dt}\right]^{\frac{1}{2}} \quad (3)$$

where the mean delay μ is given by

$$\mu = \frac{\int t h^2(t)\,dt}{\int h^2(t)\,dt}, \quad (4)$$

and the limits of integration in (3) and (4) extend over all time. We emphasize that since $h(t)$ is fixed for a given configuration, so is the delay spread. Fig. 6 presents the r.m.s. delay spread versus horizontal separation for our measured channels. In the absence of shadowing, LOS channels, whose impulse response is dominated by a short initial pulse, generally yield the smallest delay spreads, ranging from a measurement-limited 1.3 ns to about 12 ns. Unshadowed diffuse channels exhibit delay spreads that lie in the same range, but which are systematically slightly larger, due to the finite temporal spread of the dominant reflection from the ceiling. Shadowing increases the delay spread of both LOS and diffuse channels but, as might be expected, the increase is relatively modest for the latter channels. Shadowed LOS channels consistently exhibit the largest delay spreads, typically between 7 and 13 ns. In the section on performance of modulation techniques that follows, we will see that the delay spread is a reasonably accurate predictor of the multipath power penalty incurred in links using baseband on-off keying (OOK).

We have performed detailed simulations of nondirected infrared propagation [16], using a recursive technique that can account for an arbitrary number of diffuse reflections from room surfaces. For both LOS and diffuse channels, simulations are in good quantitative agreement with measured channel responses, leading us to believe that multipath infrared propagation can be well-described by simple models.

Having discussed the characteristics of nondirected infrared channels, we will now examine how to design reliable, high-speed links. We will first describe strategies for achieving a high SNR, as they are applicable to all links, independent of bit rate and modulation format. Then, we will proceed to discuss various modulation techniques, addressing their power efficiency and, for high bit rates, their robustness against multipath distortion.

Achieving a High Signal-to-Noise Ratio

The wavelength band between about 780 and 950 nm is presently the best choice for nondirected links, due to the availability of low-cost light-emitting diodes (LEDs) and laser diodes (LDs), and because it coincides with the peak responsivity of inexpensive, low-capacitance silicon photodiodes. LEDs are currently used in all commercial systems, due to their extremely low cost and because most LEDs emit light from a sufficiently large surface area that they are generally considered eye-safe. Potential drawbacks of LEDs include:

- Typically poor electro-optic power conversion efficiency of 10 to 20 percent (though new devices have efficiencies as high as 40 percent).
- Modulation bandwidth that is typically limited to tens of MHz.
- Broad spectral width (typically 25 to 100 nm), which requires the use of a wide receiver optical passband, leading to poor rejection of ambient light.
- The fact that wide modulation bandwidth is usually obtained at the expense of reduced electro-optic conversion efficiency.

LDs are much more expensive than LEDs, but offer many nearly ideal characteristics:

- Electro-optic conversion efficiencies of 30 to 70 percent.
- Wide modulation bandwidths, which range from hundreds of MHz to more than 10 GHz
- Very narrow spectrum (spectral widths ranging from several nm to well below 1 nm are available).

To achieve eye safety with a LD requires that one pass the laser output through some element that destroys its spatial coherence and spreads the radiation over an extended emitting aperture. New eye-safety regulations are likely to restrict diffuse power densities at wavelengths near 850 nm to levels of about 370 W/ m^2 for continuous viewing [17, 18]. With an LD, eye safety can be achieved using a transmissive diffuser, such as a thin plate of translucent plastic. While such diffusers can achieve efficiencies of about 70 percent, they typically yield a Lambertian radiation pattern, offering the designer little freedom to tailor the source radiation

pattern. Computer-generated holograms (CGHs) offer a means to generate arbitrary radiation patterns with efficiencies approaching 100 percent, but must be fabricated with extreme care to ensure that any residual image of the LD emission aperture is tolerably weak [17].

Infrared links are subject to two different classes of ambient light. Sunlight, skylight, and incandescent lights represent broadband visible and infrared sources that vary slowly with time. These sources induce a white, nearly Gaussian shot noise whose electrical power spectral density (PSD) is proportional to the total detected optical power. This shot noise can thus be limited by use of an optical bandpass filter or longpass filter that rejects most of the background light, but passes the desired signal. The emission of fluorescent lights, on the other hand, is intensity-modulated, either at the power-line frequency or, with new compact-ballast fixtures, at frequencies of tens of kHz. This near-d.c. interference includes harmonics that are particularly strong below 50 kHz, and that can be detected at frequencies as high as 1 MHz. While the optical emission from fluorescent lights is strongest in the visible and near-ultraviolet region, significant emission also occurs in the near-infrared, especially at 1017 nm, a spectral line of mercury. Interference due to fluorescent lighting may be controlled by a judicious combination of optical and/or electrical filtering. Longpass optical filtering reduces fluorescent-light interference, while narrow bandpass optical filtering at a carefully chosen near-infrared wavelength provides an even greater reduction of interference. Nonetheless, because of residual interference, it is usually necessary for the link to employ a passband modulation scheme that avoids the near-d.c. interference, such as PPM or multiple-subcarrier modulation (MSM). Alternatively, the link can utilize a baseband modulation technique, such as on-off keying (OOK), in conjunction with a highpass electrical filter that removes the interference. In this case, line coding or active baseline restoration is needed to prevent baseline wander. These modulation techniques will be discussed in the following section.

We will now discuss how to design the receiver optics in order to achieve a high SNR in the face of steady background illumination. A well-designed receiver will be shot-noise-limited under conditions of bright illumination, so we will assume that background-induced shot noise dominates. Under these conditions, it is appropriate to utilize a simple positive-intrinsic-negative (p-i-n) photodiode [19], rather than the avalanche photodiodes that are used in IM/DD receivers when background illumination is weak [20]. It is desirable to use a large photodetector area; we will see that the shot-noise-limited SNR is proportional to the detector area A. However, large-area detectors have high capacitance, which can limit receiver bandwidth and greatly increase receiver thermal noise [20]. Accordingly, it is desirable to reduce the required physical detector area by use of an optical concentrator, which accepts light from a large collection area and concentrates it to the somewhat smaller detector area. One type of concentrator, the dielectric compound parabolic concentrator (CPC), has been used in prototype free-space infrared links [21, 17]. The CPC is an angle-transforming device, which can concentrate light from a large input area A_{in} down to a smaller detector area A_{det}, yielding an optical gain of A_{in}/A_{det}. In doing so, however, it reduces the receiver solid angle of acceptance by the ratio $\Omega_{in}/\Omega_{det} \approx n^2 A_{det}/A_{in}$, where n is the CPC refractive index. Receivers using a single CPC-based element may be excessively directional for many applications of non-directed infrared links. An array of differently oriented CPC elements may be used to construct an angle-diversity receiver [15, 21], which can use selection or combining techniques to improve SNR or reduce the impact of multipath distortion. A second type of optical concentrator is the dielectric truncated spherical concentrator, of which a special case is the dielectric hemisphere [15, 22]. A dielectric hemisphere of index n can provide an optical gain of n^2 that is nearly omnidirectional, as long as the hemisphere radius R is at least n^2 times the detector radius r, and the two are suitably index-matched. For many applications, it is desirable to build a simple receiver containing a single, omnidirectional element, and we will restrict our attention to receivers that utilize a hemispherical concentrator.

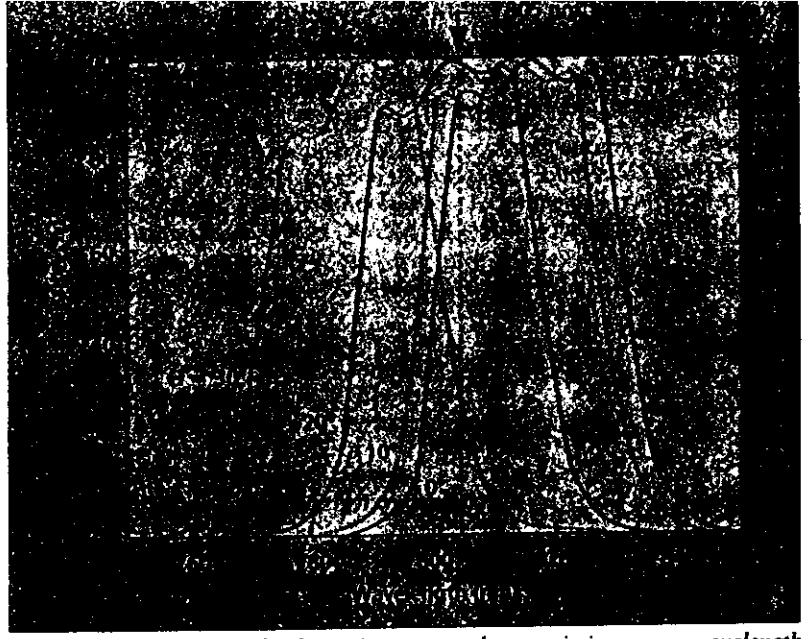


Figure 7. *Theoreteical polarization-averaged transmission versus wavelength for a 29-layer, thin-film optical bandpass filter. At normal incidence, $\theta = 0°$, the filter has a half-power bandwidth of 30 nm and a center wavelength of 816.5 nm. At non-normal incidence, the passband shifts to shorter wavelengths.*

The best rejection of ambient light can be achieved using a spectrally narrow source, such as an LD, in conjunction with a narrow optical bandpass filter. Efficient, narrowband optical interference filters can be constructed using multiple layers of thin dielectric films [23]. Unfortunately, the phase shift through the dielectric layers changes with angle of incidence, so that for non-normal incidence, the filter passband shifts to shorter wavelengths. Fig. 7 illustrates the angle-dependent passband of a typical 29-layer filter having a 30-nm half-power bandwidth at normal incidence. It is seen that the passbands for $\theta = 0°$ and $\theta = 30°$ barely overlap.

Parameter	Hemispherical filter	Planar filter
Required average transmitter power	224 mW	492 mW

Table 2. *Optimized filter bandwidth and required average transmitter power in 100 Mb/s OOK links. LOS source is optimized to cover a 5 m x 5 m room using minimum total power. A non-distorting channel, bright skylight [p_{bg} = 5.8 μW/(cm²-nm)] and BER = 10^{-9} (SNR_p = 144) are assumed. Other parameters include: detector area A = 1 cm², concentrator gain G = 4.6 dB, wavelength λ = 810 nm, and responsivity r = 0.53 A/W (from [22]).*

Fig. 8 displays two different ways that a thin-film interference filter can be combined with a hemispherical concentrator and large-area detector. As shown in Fig. 8(a), it is straightforward to place the filter between the hemisphere and detector. It has been determined through ray tracing that in this geometry, the angle-dependent filter passband makes it impossible to simultaneously achieve the desired properties of narrow optical bandwidth and wide FOV [22]. Alternatively, an optical bandpass filter can be deposited or bonded onto the outer surface of the hemispherical concentrator, as shown in Fig. 8(b). Ray tracing calculations [22] show that regardless of the angle ψ from which the signal is received, rays that reach the detector are incident upon the filter at angles θ between 0° and θ_{max}, where

$$\theta_{max} = \sin^{-1}(n/R)$$

Typically, θ_{max} is less than 30°. Thus, with a hemispherical filter, it is possible to simultaneously obtain a narrow bandwidth and wide FOV.

With the aim of optimizing the receiver optical design, we will calculate the SNR of a shot-noise-limited receiver, explicitly considering the effects of the concentrator gain, bandpass filter properties, and detector area. We consider a link using OOK modulation over a channel that is free of multipath distortion. In the following section, the performance of OOK on the ideal channel will be used as a reference for evaluating the power efficiency of OOK and other modulation techniques on multipath channels. We assume that the OOK transmission, at bit rate $1/T$, consists of rectangular pulses of duration T having a peak power of either 0 or P_p. Assuming that ones and zeros are equally probable, the average power is $P_{avg} = P_p/2$. The channel, being ideal, has a gain $H(0)$ at all frequencies. The receiver employs a concentrator with optical gain G and an optical filter that transmits the signal with efficiency η_{sig}, which represents the average over rays striking the filter at different angles with respect to its surface normal. We emphasize that η_{sig} depends on the signal angle of incidence ψ. The optical signal is detected by a *p-i-n* photodiode of area A and responsivity r. The detected current is passed through a filter having rectangular impulse response of duration T and peak value $1/\sqrt{T}$, and the filter output is sampled. The resulting samples have values of 0 or $P_p H(0) G \eta_{sig} A r \sqrt{T}$. Incident upon the receiver is a background flux of density p_{bg}, which has units of (W/cm²-nm), and is assumed to be constant over the filter passband. The filter has peak transmission η_{bg} and optical noise bandwidth Δλ. Including the concentrator gain G, a background irradiance $p_{bg} G \eta_{bg} \Delta\lambda$ is incident on the detector. This induces a shot-noise current of two-sided PSD $e p_{bg} G \eta_{bg} \Delta\lambda A r$, where e is the electronic charge, resulting in a sampled filter output with variance equal to this PSD. The peak receiver SNR is thus:

$$SNR_p = \frac{P_p^2 H^2(0) G \eta_{sig}^2 A r T}{e p_{bg} \eta_{bg} \Delta\lambda}, \quad (5)$$

and the bit-error rate (BER) is given by

$$\mathrm{BER} = Q\left(\frac{\sqrt{SNR_p}}{2}\right), \quad (6)$$

where $Q(\bullet)$ is the Gaussian Q function [24]. For example, the peak SNR should have a value of 144 (21.6 dB) to achieve a BER of 10^{-9}.

Examining (5), we see that the SNR is improved by utilization of a detector having a large area A and

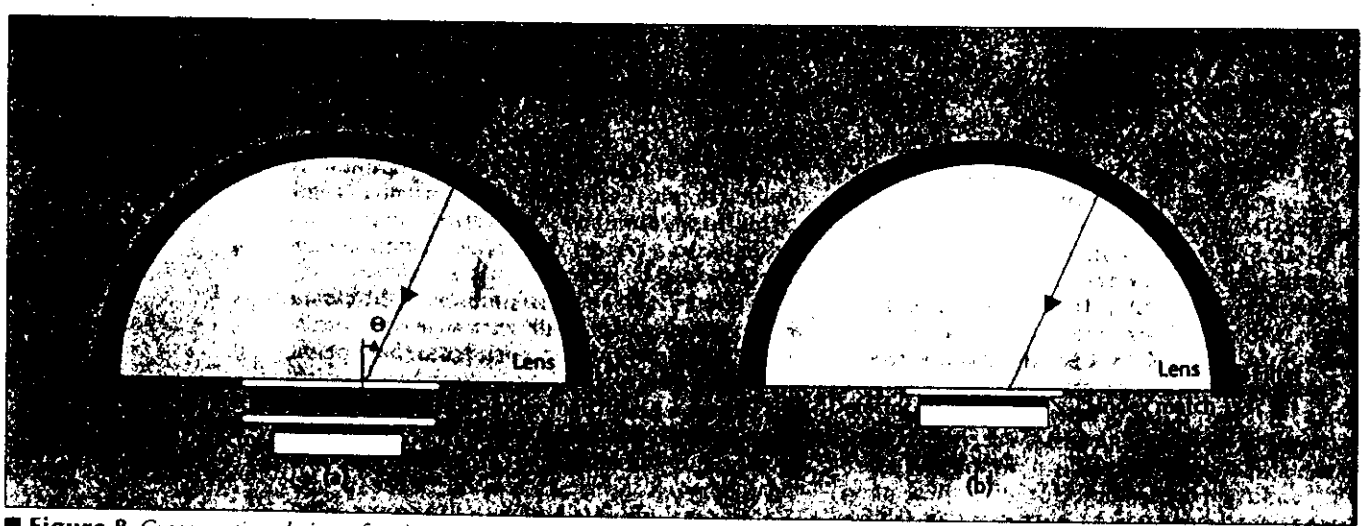


Figure 8. *Cross-sectional view of reciever structures that attempt to achieve wide field of view and narrow passband. The hemispherical concentrator of refractive index n provides a nearly omnidirectional optical gain of approximately n², as long as its radius exceeds n² times the detector radius. Structure (a) uses a flat optical bandpass filter; rays arriving at angle ψ strike the filter at angles θ that are close to ψ. Structure (b) uses a hemispherical filter; independent of the direction ψ from which light is recieved, all rays reaching the photodiode strike the filter at angles θ that are near normal incidence — typically between 0° and 30° (from [22]).*

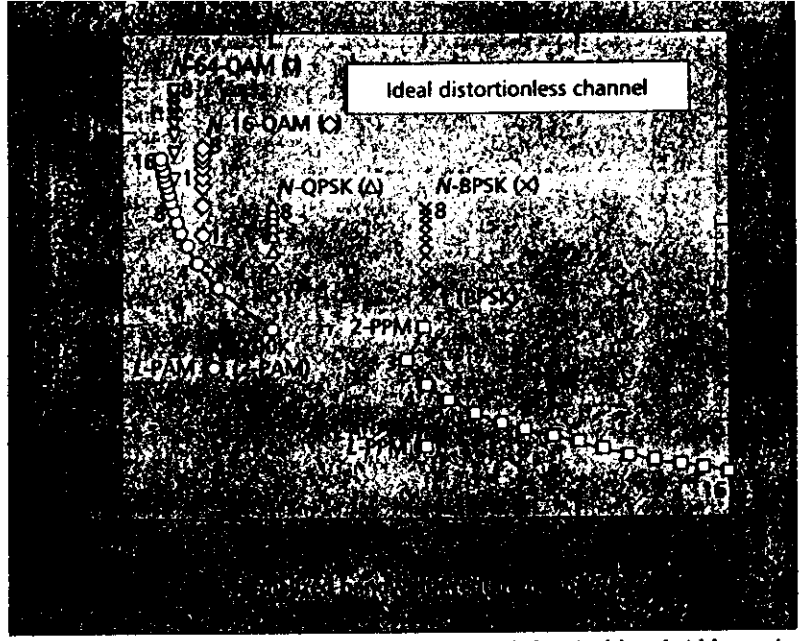
70

the highest available responsivity r, in conjunction with a concentrator having high optical gain G. The optical filter should be designed so as to maximize the figure of merit $\eta_{sig}^2/\eta_{bg}\Delta\lambda$. To first order, this implies that one should choose the narrowest possible filter bandwidth $\Delta\lambda$ that still achieves high signal efficiency η_{sig}. Assuming a hemispherical filter, as in Fig. 8(b), this can be achieved approximately as follows. One should choose the filter center wavelength so that for $\theta = 0$, the short-wavelength edge of the filter passband lies at the signal wavelength. The filter bandwidth should then be chosen so that for $\theta = \theta_{max}$, the long-wavelength edge of the passband is at the signal wavelength. For example, the filter of Fig. 7 has been designed using this procedure for a signal wavelength of 805 nm and for $\theta_{max} = 28°$. In fact, the shot-noise-limited SNR is maximized by choice of an even narrower filter bandwidth, which attenuates the signal somewhat but rejects more ambient light. We have simultaneously optimized the filter characteristics and transmitter radiation pattern of a LOS link to minimize the total transmitted power while yielding the required SNR everywhere within the transmitter coverage area [22]. Table 2 presents results for a 100 Mb/s OOK link operating over a 5 m x 5 m area. The optimized hemispherical-filter bandwidth is 11.6 nm, and the required average transmitter power is 224 mW. When a flat filter is utilized instead, the optimum filter bandwidth is increased to 71.7 nm, requiring 492 mW of transmitter power, a 3.4-dB increase.

Receiver preamplifier design is a crucial consideration for high-speed infrared links, and is discussed in detail in [22]. Among candidate preamplifier designs [20], the transimpedance amplifier probably is the best choice for most applications, due to its superior dynamic range. The preamplifier should be designed so that, under conditions of the brightest ambient illumination to be encountered, the receiver achieves sufficient SNR and is shot-noise limited, or nearly so. One should choose a *p-i-n* photodiode having low capacitance per unit area, to minimize a strong component of receiver noise whose PSD is proportional to the square of receiver input capacitance and to the square of frequency. Unfortunately, low capacitance per unit area requires that the photodiode be relatively thick, and this may lead to a transit-time limitation of receiver bandwidth [19], especially for devices illuminated through the n contact. It is also important to choose a photodiode having low series resistance. Analysis of FET-based preamplifiers [22] shows that the front-end FET should be chosen to have high transconductance g_m; the FET cutoff frequency f_T is important mainly in that the FET gate capacitance should be smaller than the photodiode capacitance.

Performance of Modulation Techniques

Non-directed infrared channels can be modeled as fixed, linear systems with additive, white, Gaussian noise, as summarized in (1). As it represents instantaneous optical power, the channel input $X(t)$ must be nonnegative, and it is related to the average transmitted optical power by (2). The channel, described by impulse response $h(t)$, exhibits multipath distortion that can induce intersymbol interference (ISI) in high-bit-rate links. When the ISI is relatively mild, it leads to an optical power penalty, but if it is severe, it may lead to a BER floor.



■ **Figure 9.** *Theoretical optical average-power and electrical-bandwidth requirements of several intensity-modulation/direct-detection schemes, in the presence of additive, white Gaussian noise, assuming an ideal, nondispersive channel. The power and bandwidth requirements are normalized to those of baseband on-off keying (from [22]).*

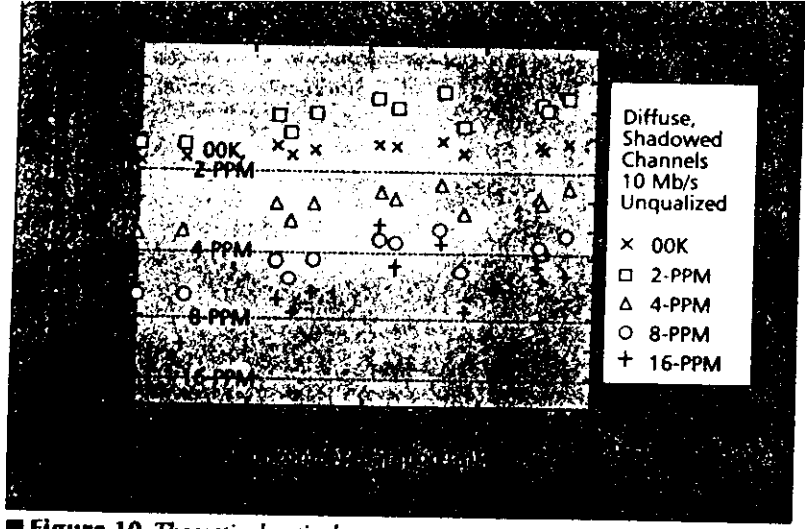
In evaluating candidate modulation techniques, the most important criterion is the optical average-power requirement, as it generally corresponds to transmitter electrical power consumption and ocular hazard. At high bit rates, one must consider the effect of multipath ISI on this power requirement, as well as any reduction of the ISI that can be achieved through techniques such as adaptive equalization. The second most important attribute is the receiver electrical bandwidth requirement,[3] as it can be difficult to achieve flat frequency response and low noise over a wide bandwidth using large-area photodiodes. Other important criteria for comparison of modulation techniques are the complexity and power consumption of a portable receiver.

Fig. 9 presents a comparison of the average power and bandwidth requirements [22] of a number of modulation schemes, assuming a distortionless channel. The power and bandwidth requirements are normalized to those of baseband OOK. These modulation techniques include single-carrier schemes, such as L-PPM and L-pulse-amplitude modulation (L-PAM), of which an important special case is 2-PAM (OOK). They also include MSM schemes that employ N subcarriers modulated using binary or quadrature phase-shift keying (BPSK or QPSK), and 16- or 64-quadrature-amplitude modulation (QAM). In the following section, we discuss the performance of several of these modulation techniques, emphasizing their power requirement in the face of multipath ISI, which we have evaluated using a large set of measured channel impulse responses. Our reference point for comparison will be the power requirement of OOK on an ideal channel, which was calculated in the previous section.

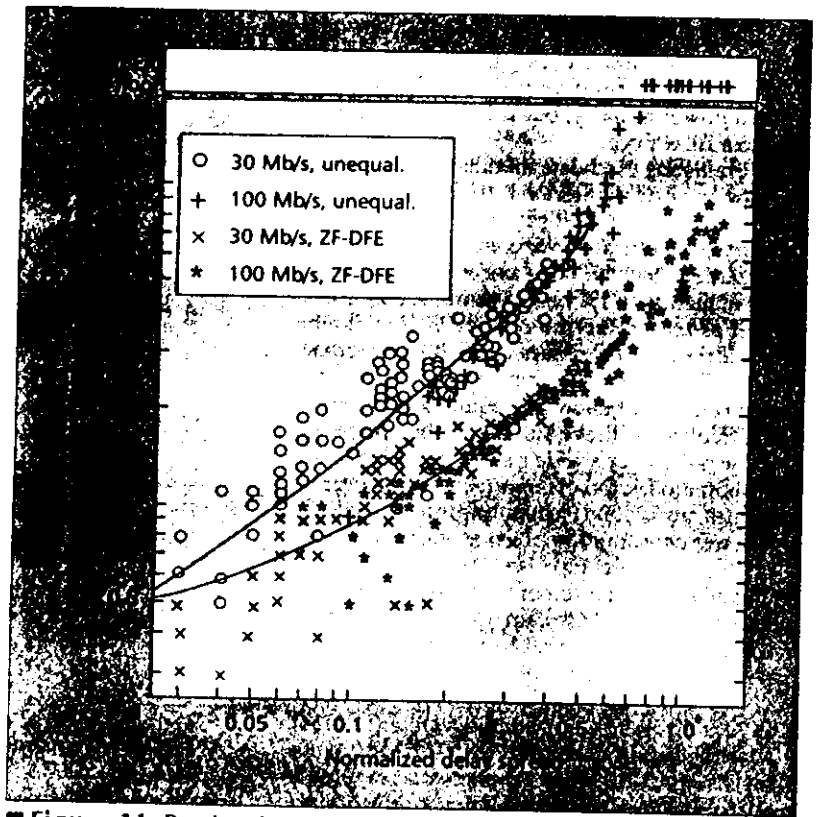
It is evident in Fig. 9 that OOK represents a

[3] *We note that the electrical bandwidth requirement of a modulation technique has little bearing on the optical bandwidth occupied by an IM signal. This optical bandwidth is dominated by the large spectral spread of practical infrared sources. For example, a 1-nm width corresponds to 469 GHz, assuming a wavelength of 800 nm.*

71



■ **Figure 10.** *Theoretical optical average-power requirement vs. horizontal separation between transmitter and receiver for five different modulation schemes at 10 Mb/s. Symbols represent penalties calculated using measured impulse responses of diffuse, shadowed channels. In these calculations, transmitter employs rectangular pulses, of duration T for OOK, and of duration T/L for L-PPM; receiver employs a filter matched to the transmitter pulse shape, and BER is 10^{-6}. Dashed lines represent performance over an ideal channel. The reference (0 dB) is for an OOK link on a distortionless channel (from [25]).*



■ **Figure 11.** *Baseband on-off keying at 30 and 100 Mb/s, unequalized and with zero-forcing decision-feedback equalization: optical average-power requirement versus normalized delay spread (r.m.s. delay spread divided by bit duration T). The reference (0 dB) is for an OOK link on a distortionless channel. Symbols represent power requirements calculated using measured impulse responses of four channel types (LOS and diffuse, both unshadowed and shadowed). In these calculations, transmitters utilize rectangular pulses of duration T and BER is 10^{-9}. Unequalized and equalized receivers employ five-pole Bessel filters with 3-dB cutoffs of 0.6/T and 0.45/T, respectively. Curves represent a least-squares linear fit of the power penalty (in dB) to the normalized delay spread, for normalized delay spreads less than 0.5 (from [12]).*

good compromise between power requirement and bandwidth requirement, at least on the ideal, distortionless channel. When considering the impact of multipath ISI, the power requirement of OOK links can be calculated using the technique described in [16]. The impulse responses of non-directed infrared channels we have characterized [12] contain significant energy only within the first 70 ns after their first non-zero excursion. It is thus not surprising that at a bit rate of 10 Mb/s, where the bit duration is 100 ns, ISI induces little impairment. For example, normalized power requirements of 10-Mb/s OOK links on shadowed, diffuse channels do not exceed 1.0 dB, as indicated by the x symbols in Fig. 10. This figure also displays the power requirement of links using several different types of L-PPM, which will be discussed below. Clearly, unequalized OOK transmission is feasible at 10 Mb/s. At higher bit rates, OOK links incur more significant degradation from multipath ISI [12]. For example, at 30 Mb/s, normalized power requirements on shadowed, diffuse channels do not exceed 4.3 dB. While relatively high, these power requirements may be acceptable for some applications, particularly when transmitting from a base station. At a bit rate of 100 Mb/s, however, very large power requirements or even BER floors are incurred on all four types of channels (LOS and diffuse, with and without shadowing), suggesting that unequalized OOK is not practical in this case. Fig. 11 displays the normalized power requirements of unequalized OOK links transmitting on all four channel types at bit rates of 30 and 100 Mb/s, denoted by o and + symbols, respectively. It is seen that there is a systematic relationship between the normalized power requirement and the normalized channel delay spread, which is the channel delay spread (3) divided by the bit duration T. Despite the fact that the various impulse responses differ in their detailed features, for purpose of calculating the impact of ISI on OOK links, channels of all four types can be roughly described by a single-parameter model.

A practical, though sub-optimal, means to reduce the multipath ISI penalty is by using a decision-feedback equalizer (DFE) that can adapt automatically to the channel impulse response [24]. Considering measured channel responses, the power requirements of OOK links using DFEs have been evaluated in [12]. At 30 Mb/s, normalized power requirements for all channels are less than 2.3 dB. Thus, OOK with a DFE is a feasible technique for transmission at 30 Mb/s. At 100 Mb/s, no BER floors are observed with a DFE, in contrast to the unequalized case. With a DFE, power requirements for unshadowed LOS and diffuse channels are less than 6.7 and 5.7 dB, respectively, while those for shadowed LOS and diffuse channels do not exceed 9.1 and 7.1 dB, respectively. These power requirements, while high, may be small enough to make OOK with a DFE practical at 100 Mb/s, particularly for transmission from a base station. Fig. 11 presents the normalized power requirements of OOK links with a DFE operating at 30 and 100 Mb/s, denoted by x and • symbols, respectively, transmitting over all four channel types. Again, we observe a systematic relationship between the normalized power requirement and normalized delay spread.

While we have considered zero-forcing DFEs to simplify calculation of the power requirements, their performance at low BER is virtually identical to that

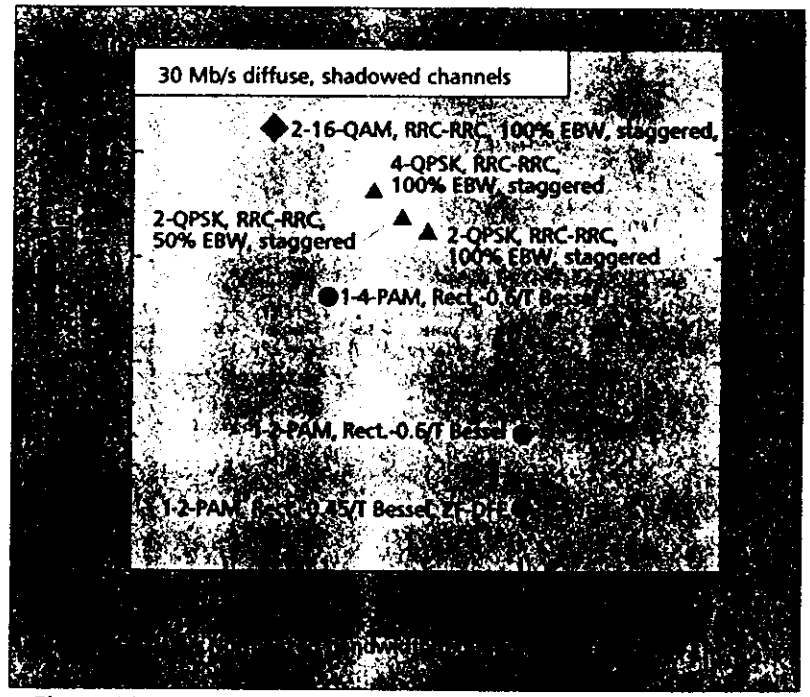
of DFEs adapted according to the minimum-mean-square error criterion [24, 25]. In practice, a DFE typically would be implemented using digital or discrete-time analog signal-processing techniques. Numerical simulation [25] of 100 Mb/s OOK with constrained-complexity DFEs have shown that little degradation is caused by either limiting the forward and reverse filters to three $T/2$-spaced taps and five T-spaced taps, respectively, or by employing four-bit quantization of tap weights. Adaptation to a training sequence using the least-mean-squares algorithm is found to occur within about 200 bits, i.e., about 2 μs. Most indoor applications of nondirected infrared links will not involve rapidly moving receivers, so that channel impulse responses will change significantly only on time scales of tens to hundreds of ms. Thus, we expect that once it is initially adapted to the channel response, a DFE should be able to track easily any changes in that response.

As mentioned above, fluorescent lighting may induce near-d.c. interference in wireless infrared receivers. The impact of this interference on a baseband OOK receiver may be reduced by electrical highpass filtering of the preamplifier output. To avoid excessive baseline wander in a baseband OOK link, however, the highpass filter cut-on frequency cannot be higher than about 10^{-3} of the bit rate [26], unless line coding or active baseline restoration is utilized [25].

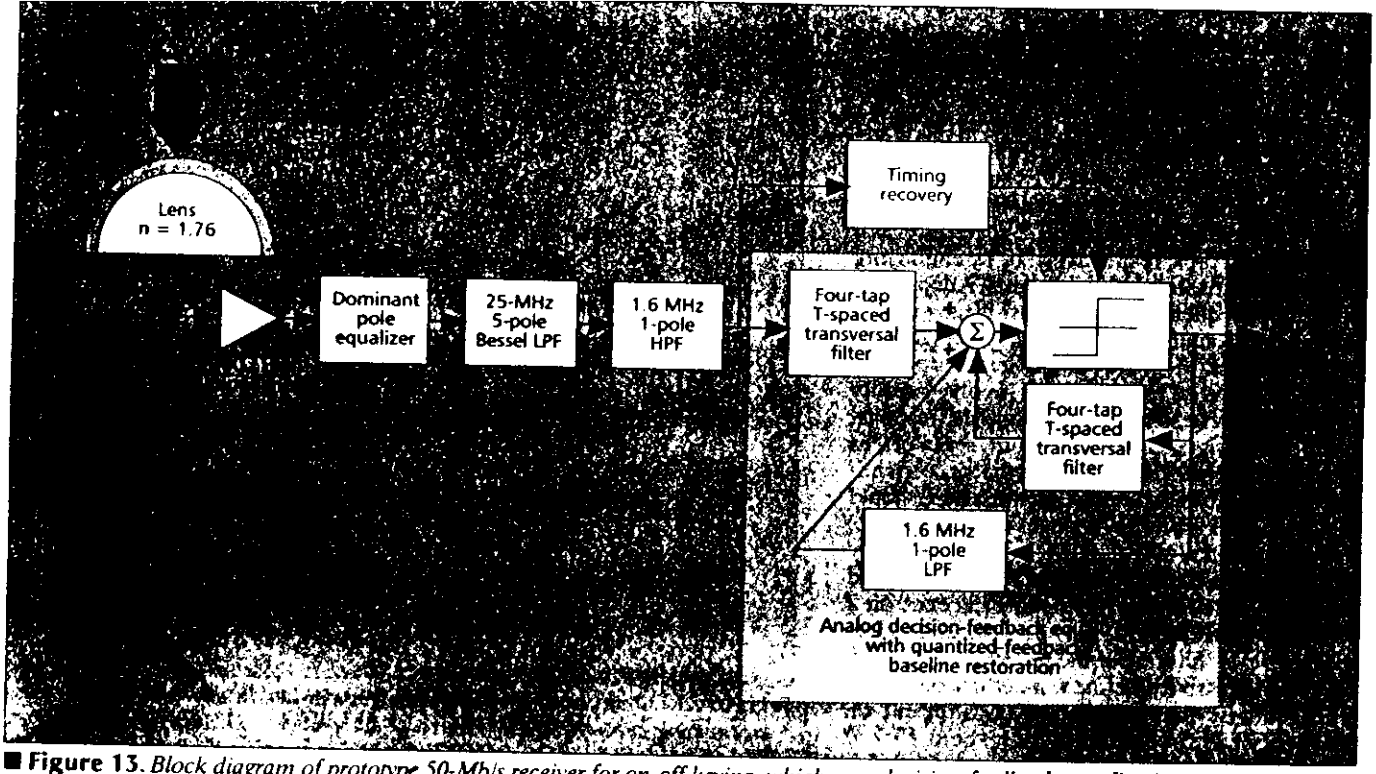
Considering L-PAM in general, as L is increased from 2 (recall that OOK is equivalent to 2-PAM), there is a monotonic decrease in bandwidth requirement, at the expense of a monotonic increase in power requirement on an ideal channel (Fig. 9). For a given bit rate, as L increases and the symbol duration increases, the decrease in noise admitted by the corresponding receiver filter is not sufficient to overcome the reduction in noise immunity associated with closer spacing of signal levels, as in conventional systems [24]. Considering multipath channels, one might hope that as L increases and, for a given bit rate, the symbol duration increases relative to the channel delay spread, a reduction in ISI penalty might overcome the poor inherent power efficiency of higher-level L-PAM, making it more efficient than OOK. Unfortunately, this is not the case. For example, at 30 Mb/s, unequalized 4-PAM has a normalized power requirement of 5.3 dB, compared to 2.7 dB for unequalized 2-PAM (OOK), averaged over 17 diffuse, shadowed channels. These two power requirements are compared in Fig. 12, where they are indicated by • symbols. This figure also presents the power requirements of several MSM schemes, which will be discussed below. We conclude that OOK is probably the best L-PAM technique for most applications.

L-PPM is a transmission technique [22, 26] that offers an improvement in average-power efficiency over OOK, at the expense of an increased bandwidth requirement (Fig. 9). This technique utilizes symbols consisting of L time slots, which we will refer to as chips. A constant power LP_{avg} is transmitted during one of these chips (P_{avg} is the average transmitted optical power), and zero power is transmitted during the remaining $L-1$ chips, thereby encoding $\log_2 L$ bits in the position of the non-zero chip. For a given bit rate, L-PPM requires more bandwidth than OOK by a factor $L/\log_2 L$, e.g., 16-PPM requires four times more bandwidth than OOK. On distortionless channels, L-PPM yields a decrease in average-power requirement that improves steadily with increasing L; the increased noise associated with a $(L/\log_2 L)$-fold wider receiver noise bandwidth is outweighed by the improved noise immunity arising from a L-fold increase in peak power. The excellent average-power efficiency of L-PPM can result in a significant decrease in transmitter power consumption, making the technique especially useful for portable transmitters. Because the power spectra of L-PPM vanishes at d.c. for all values of L, it is possible to pass a received L-PPM waveform through a highpass filter having a cut-on frequency as high as 0.05 to 0.10 times the bit rate with little baseline wander [26]. This makes L-PPM an excellent choice in the presence of near-d.c. interference from fluorescent lighting. Two drawbacks of L-PPM as compared to OOK should be noted, i.e., an increased transmitter peak-power requirement, and the need for more precise synchronization.

When L-PPM is transmitted over multipath channels, the non-zero transmitted chips can induce interference in chips both within the same symbol (intrasymbol interference) and in adjacent transmitted symbols (intersymbol interference); we will refer to these effects collectively as ISI. For transmission at a fixed bit rate on a given channel, the impact of ISI should generally increase steadily with increasing L, due to the decrease in duration of the L chips within each symbol. As L is increased, the ISI penalty will eventually overcome the inherent average-power efficiency associated with



■ **Figure 12.** *Optical average-power and electrical-bandwidth requirements of several multiple-subcarrier (MSM) and pulse-apmplitude modulation (PAM) schemes. The reference (0 dB) is for an OOK link on a distortionless channel. Power requirements represent the average of values calculated using the measured impulse responses of 17 diffuse, shadowed channels. MSM schemes considered include two staggered 16-QAM subcarriers and two and four staggered QPSK subcarrieres, with root-raised-cosine transmit and receive filters. PAM schemes include 2-PAM (OOK) and 4-PAM with rectangular-pulse transmit filters and five-pole Bessel receive filters; 2-PAM with zero-forcing decision-feedback equalization is also considered (from [28]).*



■ **Figure 13.** *Block diagram of prototype 50-Mb/s receiver for on-off keying, which uses decision-feedback equalization.*

large L. We have analyzed the impact of ISI on this transmission technique, and have evaluated the power requirement of unequalized L-PPM links using measured channel impulse responses [26]. The average-power requirement of OOK and several L-PPM orders are compared in Fig. 10, for transmission at 10 Mb/s over diffuse, shadowed channels. At this bit rate, while L-PPM links incur ISI penalties that increase with increasing L, 16-PPM still yields superior average-power efficiency on all four types of channels (LOS and diffuse, with and without shadowing). At the higher bit rate of 30 Mb/s, the same general trends are evident. However, the optical power requirements are higher, and the power requirement does not decrease steadily with increasing L for all channels. For the shadowed LOS configurations, higher PPM orders $L \geq 4$ incur BER floors on several channels, demonstrating once again that for applications where shadowing is probable, a diffuse system is more robust. At a bit rate of 100 Mb/s, excessive or infinite power penalties are incurred on all channels except unshadowed LOS channels, where 8-PPM is consistently found to yield the best power efficiency.

Given its inherent power efficiency, it would be desirable to develop techniques to mitigate the impact of multipath ISI on L-PPM, in the hope that this might permit reliable transmission at very high bit rates. Recently, Barry has discussed maximum-likelihood sequence detection of L-PPM in the presence of ISI [27], and has shown that it can be implemented by generalization of the Viterbi algorithm to the case of a vector channel. He has also discussed several sub-optimal adaptive equalization techniques, which include linear and decision-feedback equalizers operating at either the chip or symbol rates, as well as hybrid DFEs that make use of tentative chip decisions to cancel intrasymbol interference, but use more reliable symbol decisions to cancel intersymbol interference. We are currently using measured channel characteristics to quantify the performance of these detection and equalization schemes.

MSM is a technique that promises the flexibility of frequency-division multiplexing and multi-level modulation, while maintaining the simplicity of IM/DD [22, 28]. In MSM, several independent bit streams are modulated onto carriers at several frequencies (say, of the order of 1 to 100 MHz). The modulated subcarriers are summed together to form a frequency-division multiplexed signal, and this signal is used to modulate the intensity of an optical transmitter. After transmission and optical-to-electrical conversion, the individual bit streams can be recovered using multiple bandpass demodulators. Through simultaneous transmission of several narrow-band subcarriers, MSM may make possible very high aggregate bit rates without requiring adaptive equalization to overcome ISI, and may allow individual receivers to process only a subset of the total transmission. While MSM is less power-efficient than OOK or L-PPM, it may be well-suited for transmission of multiplexed bit streams from a base station to a collection of several portable receivers.

The optical average-power and bandwidth requirements of several MSM schemes on ideal channels have been derived in [22], and are summarized in Fig. 9. A single BPSK or QPSK subcarrier requires 1.5 dB more optical power than OOK; BPSK and OOK transmissions of average power P_{avg} are equivalent to binary antipodal signals plus a d.c. bias P_{avg} carrying no information, but the BPSK waveform uses sinusoidal pulses having 3-dB less electrical power, thus requiring 1.5 dB more optical power for achievement of the same receiver SNR. Single 16- and 64-QAM subcarrier schemes are less power-efficient than BPSK and

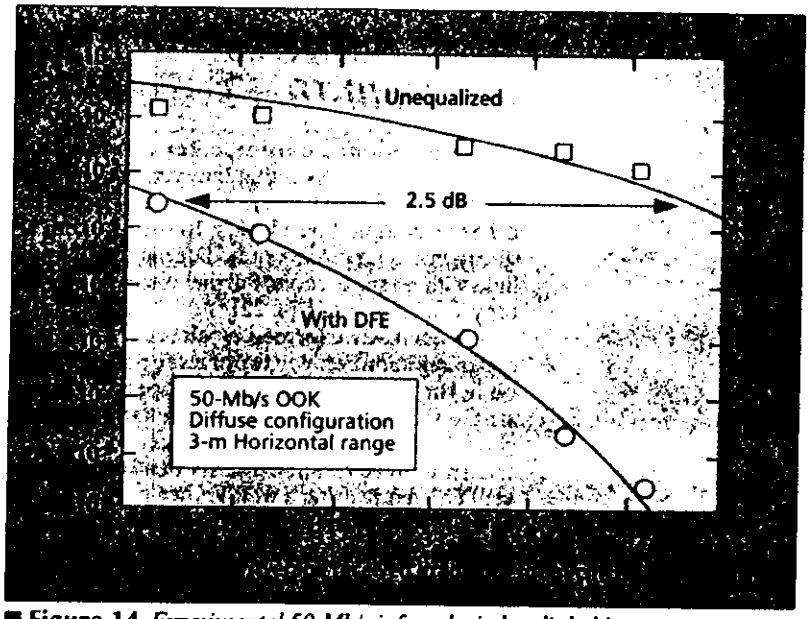
74

QPSK for conventional reasons [24]. An N-subcarrier transmission requires more average power than the corresponding single subcarrier scheme by a ratio that increases steadily with increasing N, because the amplitude of each subcarrier must not exceed P_{avg}/N, to insure that the transmitted optical-power waveform $X(t)$ is non-negative.

When MSM is transmitted over a multipath channel, several effects further degrade its power efficiency, as compared to transmission of OOK over an ideal channel [28]. As multipath channels are generally lowpass in nature, subcarriers are subject to an attenuation that generally increases with increasing subcarrier frequency. In addition, subcarriers may be subject to ISI, interference between in-phase and quadrature phases of one subcarrier, and interference between adjacent subcarriers that may overlap partially in frequency. To reduce these three interferences, it is desirable to use a large number of subcarriers, but this leads to an excessive penalty for large N, as described above. We have evaluated the performance of a large number of different MSM schemes, for transmission at total bit rates of 30 and 100 b/s over measured multipath channels [28]. The amplitudes of all subcarriers were maintained equal, and the average optical power was scaled as necessary to achieve the required BER. We found that the best performance was achieved typically by two and four-subcarrier formats. For example, Fig. 12 presents the power and bandwidth requirements of the MSM schemes that yielded the best BER-floor-free, average-power performance at 30 Mb/s. Averaged over 17 diffuse, shadowed channels, the best normalized power requirement of 6.5 dB was achieved by two QPSK subcarriers using 100 percent-excess bandwidth, root-raised-cosine pulses. While this is 5.2 dB more than the power required using OOK with a DFE, the increased power may be accommodated by a base-station transmitter. At a bit rate of 100 Mb/s, the best MSM format, 2-QPSK, yielded a normalized power requirement of 11.1 dB, averaged over the same 17 channels. Improvement in MSM performance, at the expense of a variable transmission bit rate, can be achieved through dynamic carrier selection, i.e., by not transmitting at a subcarrier frequency at which the channel attenuation and/or dispersion is excessive.

Experimental 50-Mb/s Diffuse Infrared Link

Over the past 18 months, we have been designing and constructing an experimental link to test the performance limits of high-speed non-directed infrared communication. In its initial form, this prototype uses baseband OOK transmission at 50 Mb/s [29]. Our transmitter uses a cluster of eight LDs whose output is passed through a translucent plastic diffuser to create an approximately Lambertian radiation pattern having 475-mW average power at a wavelength of 805 nm. In typical operation, the transmitter emission is directed upward toward the ceiling, creating a diffuse link configuration. A block diagram of the receiver is shown in Fig. 13. The receiver employs an optical "antenna" of the design shown in Fig. 8(b). A 1-cm^2 silicon p-i-n detector is index-matched to a hemispherical concentrator of 2-cm radius, having a refractive index of 1.76. An optical bandpass filter designed to have the wavelength-dependent transmission characteristics shown in Fig. 7 is bonded to the hemisphere's outer surface. This filter-concentrator combination achieves a bandwidth of 30 nm, a net gain of 1.5 dB, and a FOV of 65°. The photodiode capacitance of 35 pF, in conjunction with the preamplifier load resistance of 10 kΩ, leads to a 455-kHz pole that is compensated by a passive R-C circuit. This receiver achieves a 3-dB cutoff frequency of 25 MHz, which is limited by the transit time of holes across the depletion region of the photodiode, which is illuminated through the n contact. Our receiver has an equivalent input noise current density (one-sided) of 7.8 pA/$\sqrt{\text{Hz}}$, averaged over the bandwidth of the 25-MHz Bessel lowpass filter. Residual interference from fluorescent lighting is removed using a 1.6-MHz, single-pole highpass filter, and quantized feedback through a 1.6-MHz, single-pole lowpass filter is used to prevent baseline wander. In order to reduce the impact of multipath ISI, our receiver employs a DFE, with the forward and reverse filters realized using cable delays and manually adjusted, variable-gain amplifiers. Both filters have four taps; those of the forward filter are half-baud-spaced, while those of the reverse filter are baud-spaced.



■ **Figure 14.** *Experimental 50-Mb/s infrared wireless link: bit-error rate vs. average received signal irradiance for unshadowed, diffuse transmission over a 3-m horizontal range, with and without decision-feedback equalization. Equalization provides a 5-dB enhancement of signal-to-noise ratio, reducing the optical power requirement by 2.5 dB. No background light is present. The curves represent best-fit Q-functions.*

Fig. 14 presents measurements of the BER achieved by a diffuse link transmitting over a 3-m horizontal range, with no ambient light present. It is seen that the DFE is extremely effective in mitigating multipath ISI, reducing the optical power penalty by 2.5 dB (equivalent to an SNR improvement of 5 dB). In the absence of illumination, the receiver sensitivity is -32.5 dBm/cm^2 at 10^{-7} BER. Fluorescent lighting induces an optical power penalty of only 0.1 dB, while bright skylight leads to a penalty up to 1.5 dB. In the presence of bright skylight, the link achieves a horizontal range of 2.9 m at 10^{-7} BER.

Interference in Multi-User Systems

As mentioned earlier, a key advantage of infrared LANs is that interference can arise only between users not separated from each other by opaque boundaries. In this section, we discuss briefly how to accommodate multiple infrared links within a single room. Further discussion on this topic can be found in [4, 22].

Ad hoc LANs, formed by direct peer-to-peer communication between two or more potables located in the same room, typically operate using a single shared channel, e.g., one transmission wavelength accommodating packetized, L-PPM transmission [8]. Such a LAN can be operated very efficiently using a well-known protocol, such as CSMA-CA, though protocol modification is sometimes necessary to provide for the possibility of "hidden nodes," i.e., situations in which some of the colocated portables cannot communicate with each other.

LANs that involve duplex communication between portables and base stations can employ a single transmission wavelength for uplink and downlink communications. Single-carrier modulation schemes, such as OOK and LPPM, will provide a single channel that must be shared by uplinks and downlinks using an appropriate protocol. Multiple-carrier modulation schemes, such as MSM or hybrid baseband/MSM schemes,can provide multiple, non-interfering channels for separate use by uplinks and downlinks. The single-wavelength approach has the advantage that peer-to-peer communication between portables is also possible, making use of the same portable receiver used for downlink reception. Use of separate wavelengths for uplink and downlink transmission can increase the capacity of LANs involving uplink and downlink communications. When such wavelength duplexing is used, however, portables must be able to receive at two wavelengths if peer-to-peer communication between portables is also desired.

The prospects appear good that much higher transmission speeds can be achieved, and that infrared will play a significant role in future high-capacity indoor wireless LANs.

In LANs that utilize base stations, rooms up to 10 m x 10 m in size, or even larger, may be served by a single base station, depending on uplink and downlink transmission range. In this case, downlink transmissions will not be subject to interference from other downlinks, and will be essentially free of interference if the LAN employs wavelength duplexing of uplink and downlink transmissions. Downlinks may interfere with each other in larger rooms that require multiple base stations. In this case, the greatest aggregate downlink capacity can be achieved by partitioning the room into non-overlapping regions covered by different base stations, taking advantage of the very rapid falloff of interference with distance (see below). Unfortunately, this approach will introduce "dead zones" between base-station coverage regions. In rooms that can be covered by two or three base stations, it is possible to achieve full coverage by having all base stations transmit in unison, but this may lead to an unacceptable increase of multipath distortion. The best general means of achieving full coverage of large areas is probably to adopt a cellular approach, partitioning the downlink into time intervals (with single-carrier modulation scheme) or frequency bands (with MSM schemes) that are not used in adjacent cells, but that may be reused in distant cells.

It is worthwhile to discuss briefly the modeling of interference in non-directed infrared systems. Consider M simultaneous IM transmissions $X_j(t)$, $j = 1, \ldots, M$, which are incident upon a DD receiver. Let $h_j(t)$ denote the impulse response of the channel between transmitter j and the receiver. Then the total received photocurrent $Y(t)$ is given by

$$Y(t) = \sum_{j=1}^{M} X_j(t) \otimes h_j(t) + n(t) . \qquad (7)$$

We emphasize that $Y(t)$ is linear in each of the IM envelopes $X_j(t)$, and that there is no need to consider the relative phases of the underlying optical carriers. The derivation of expression (7) is a simple generalization of the derivation of (1), which is provided in [12].

As an example of a multiuser system with interference, we will consider the case of OOK transmission. We assume that the bit rate is low enough that channel dispersion can be neglected. The desired signal, of peak power P_p, transmitted over a channel having d.c. gain $H(0)$. There are K-independent, synchronized interferers, for which the corresponding quantities are P_{pj} and $H_j(0), j = 1, \ldots, K$. We assume that ones and zeroes are equipropable in all transmissions. Using (7), it is straightforward to derive the peak SIR:

$$SIR_p = \frac{4P_p^2 H^2(0)}{\sum_{j=1}^{K} P_{pj}^2 H_j^2(0)} . \qquad (8)$$

The numerator and denominator of (8) are proportional to equivalent electric powers, as one would expect. We note, however, that the equivalent d.c. electrical power gains $H^2(0)$ and $H_j^2(0)$ are equal to the inverse squares of the respective optical path losses. For example, suppose that the path loss increases as the fourth power of distance (Fig. 5). In this case, the SIR increases proportional to the eighth power of the interferer separation, in contrast with the fourth-power increase that would seen in a system using linear detection techniques. While this advantage of wireless infrared links using IM/DD is not yet widely appreciated, it may lead to significant enhancement of infrared LAN capacity, as compared to LANs employing radio links.

Summary and Conclusions

In this article we have reviewed the advantags and drawbacks of non-directed infrared radiation as a communication medium for indoor wireless LANs. The physical characteristics of IM/DD infrared channels, including path loss and multipath distortion, were described. We discussed techniques

for achieving a high receiver SNR in the face of high path loss and intense ambient infrared radiation. Several promising modulation formats were described, including OOK, L-PPM, and MSM, and we quantified the performance of these techniques on real multipath channels. We have described preliminary tests of a 50-Mb/s diffuse infrared link, showing that OOK transmission with a DFE is feasible at this bit rate. Finally, we discussed interference in multi-user infrared systems.

At this writing, non-directed infrared LANs operating at bit rates up to 4 Mb are commercially available. The prospects appear good that much higher transmission speeds can be achieved, and that infrared will play a significant role in future high-capacity indoor wireless LANs. In order for infrared LANs to each their full potential, much R & D work remains to be performed. Fertile research topics include optimized design of transmitter and receiver optics, equalization techniques for L-PPM, diversity reception techniques, techniques to mitigate interference in multi-user systems, and low-power implementation of relevant analog and digital circuitry.

Acknowledgements

The authors are grateful to E. A. Lee and D. G. Messerschmitt for their contributions in the early phases of this research. This work has been supported by IBM, DEC, AT&T Foundation, California MICRO Program, NSF-PYI-ECS-9157089, and graduate fellowships from NSF, Canadian NSERC, and IBM.

References

[1] F. R. Gfeller and U.H. Bapst, "Wireless In-House Data Communication via Diffuse Infrared Radiation," Proceedings of the IEEE, vol. 67, no. 11, Nov. 1979, pp. 1474-1486.

[2] M. D Kotzin and A. P. van den Heuvel, "A Duplex Infra-Red Systems for In-Building Communications," Proc. IEEE Vehicular Technology Conference, 1986, pp. 179-185.

[3] T. S. Chu and M. J. Gans, "High Speed Infrared Local Wireless Communication," *IEEE Commun. Mag.*, vol. 25, no. 8, Aug. 1987, pp. 4-10.

[4] A. Lessard and M. Gerla, Wireless Communication in the Automated Factory Environment," *IEEE Network*, vol. 2, no. 3, May 1988, pp. 64-69.

[5] J. R. Barry *et al.*, "High-Speed Non-Directive Optical Communication for Wireless Networks," *IEEE Network*, vol. 5, no. 6, Nov. 1991, pp. 44-54.

[6] For example, a bit rate of 125 Mb/s over a 30-m range, using transmitter and receiver fields of view of 1° x 1° and 6° x 6°, respectively, is achieved by the JDL 889-LS transceiver manufactured by JOLT, Ltd., Jerusalem, Israel.

[7] G.M. Kaye, "Yes, SIR, That's Hewlett-Packard's Baby," *Photonics Spectra*, July 1993, p. 66.

[8] International Business Machines Corporation, Armonk, NY

[9] Photonics Corporation San Jose, CA.

[10] Spectrix Corporation, Evanston, IL.

[11] C. T. Baumgartner, Spectrix Corporation, personal communication

[12] J. M. Kahn, W. J. Krause, and J. B. Carruthers, "Experimental Characterization of Non-Directed Indoor Infrared Channels," submitted to *IEEE Trans. on Commun.*, Aug. 1993

[13] R. D. Gitlin, J. F. Hayes and S. B. Weinstein, Data Communications Principles, (New York: Plenum Press, 1992).

[14] J. D. Parsons, The Mobile Radio Propagation Channel, (New York: Halsted Press, 1992).

[15] M. D. Kotzin, "Short-Range Communications Using Diffusely Scattered Infrared Radiation," Ph.D. Dissertation, Northwestern University, 1981.

[16] J. R. Barry *et al.*, "Simulation of Multipath Impulse Response for Wireless Optical Channels," *IEEE J. Sel. Areas in Commun.*, vol. 11, no. 3, April, 1993, pp. 367-379.

[17] P. P. Smyth *et al.*, "Optical Wireless Local Area Networks-Enabling Technologies", *B T Technol J.*, vol 11, no. 2, April 1993.

[18] IEC825: "Radiation Safety of Laser Products' Equipment Classification, Requirements, and User's Guide" (1984), plus amendments 1 (1990) and 2 (1994).

[19] S. M. Sze, Physics of Semiconductor Devices, 2nd Ed., (New York: John Wiley & Sons, 1981).

[20] B. L. Kasper, "Receiver Design," in Optical Fiber Telecommunications II, S.E. Miller and I.P. Kaminow, Eds., (Academic Press, 1988), pp. 689-723.

[21] R. L. Poulin, D. R. Pauluzzi, and M. R. Walker, "A Multi-Channel Infrared Telephony Demonstration System for Public Access Applications," Proc. 1992 IEEE Conf. on Sel. Topics in Wireless Commun., Vancouver, B.C., Canada, June 25-26, 1992, pp. 286-292.

[22] J. R. Barry, Wireless Infrared Communications, (Boston, MA: Kluwer Academic Publishers, 1994).

[23] J. D. Rancourt, *Optical Thin Films*, (New York, NY: Macmillan Publishing, 1987).

[24] E. A. Lee and D. G. Messerschmitt, Digital Communication, Second Edition, (Boston, MA: Kluwer Academic Publishers, 1994).

[25] M. D. Audeh and J. M. Kahn, "Performance Evaluation of Baseband OK for Wireless Indoor Infrared LANs Operating at 100 Mb/s," to be published in *IEEE Trans. on Commun.*

[26] M. D. Audeh and J. M. Kahn, Performance Evaluation of L-Pulse Position Modulation on Non-Directed Indoor Infrared Channels, submitted to *IEEE J. Sel. Areas Commun.*, Nov. 1993.

[27] J. R. Barry, "Sequence Detection and Equalization for Pulse-Position Modulation," presented at SUPERCOMM/ICC '94, New Orleans, LA, May 1-5, 1994.

[28] J. B. Carruthers, "Performance and sign of Multiple-Subcarrier Modulation Schemes for Wireless Indoor Infrared LANs," M.S. Report, University of California at Berkeley, Dec. 1993.

[29] G. W. Marsh and J. M. Kahn, "50-Mb/s Diffuse Infrared Free-Space Link Using On-Off Keying with Decision-Feedback Equalization," submitted to *IEEE Photon. Technol. Lett.*

Biographies

JOSEPH M. KAHN [M '87] received A.B. (1981), M.A. (1983), and Ph.D. (1986) degrees from U.C. Berkeley, all in physics. From 1987 to 1990 he was a member of technical staff in the Lightwave Communications Research Department of AT&T Bell Laboratories, where he performed research on multi-gigabit-per-second coherent optical fiber transmission systems and related device and subsystem technologies. He demonstrated the first BPSK-homodyne optical fiber transmission system, and achieved world records for receiver sensitivity in multi-gigabit-per-second systems. He joined the faculty of U.C. Berkeley in 1990, and is an associate professor in the Department of Electrical Engineering and Computer Sciences, where his research interests include optical fiber communication networks and transmission systems, LANs using free-space optical links, and optical interconnects in digital systems. He is a recipient of the National Science Foundation Presidential Young Investigator Award, and is a member of the IEEE Lasers and Electro-Optics Society. He is serving currently as a technical editor of *IEEE Personal Communications Magazine*.

JOHN R. BARRY received a B.S. from the State University of New York at Buffalo in 1986, and M.S. and Ph.D. degrees from U.C. Berkeley in 1987 and 1992, respectively, all in electrical engineering. He is currently an assistant professor in the School of Electrical and Computer Engineering at the Georgia Institute of Technology, where his research interests include digital communication theory and broadband wireless communications. He received the 1992 David J. Griep Memorial Prize and the 1993 Eliahu Jury Award from U.C. Berkeley, and a 1993 IBM Faculty Development Award.

MALIK AUDEH received a B.S. in 1990 from the University of Illinois at Urbana-Champaign and an M.S. in 1992 from U.C. Berkeley, both in electrical engineering. Currently, he is pursuing a Ph.D. at U.C. Berkeley in electrical engineering, where he is continuing his earlier research on wireless indoor infrared communication. His research interests include digital communications, digital signal processing, and optical communications.

JEFFREY B. CARRUTHERS [S '88] received a B. Eng. degree in computer systems engineering from Carleton University in 1990, and joined Bell-Northern Research in Ottawa, working on development of SONET-compliant products until 1991. In 1993 he received the M.S. degree in electrical engineering from U.C. Berkeley, where he is currently working toward a Ph.D. degree. His research interests include diversity reception and modulation techniques for wireless optical communication.

WILLIAM J. KRAUSE received a B.S. in electrical engineering from Johns Hopkins University, in 1990. In 1992 he received an M.S. in electrical engineering from U.C. Berkeley. His M.S. research involved wideband characterization of multipath free-space infrared channels. He is a member of Eta Kappa Nu and Tau Beta Pi.

GENE MARSH received a B.S. in computer science and electrical engineering from Rose-Hulman Institute of Technology in 1985, and an M.S. in electrical engineering from Stanford University in 1987. He is currently working on a Ph.D. in electrical engineering at U.C. Berkeley, specializing in wireless optical communications.

An evolution toward three large groups of applications and services

Wireless Personal Communications: What Is It?

DONALD C. COX

Wireless Personal Communications has captured the attention of the media, and with it, the imagination of the public. Hardly a week goes by without one seeing an article on the subject appearing in a popular U.S. newspaper or magazine. Articles ranging from a short paragraph to many pages regularly appear in local newspapers, as well as in nationwide print media, e.g., *The Wall Street Journal*, *The New York Times*, *Business Week*, and *U.S. News and World Report*. Countless marketing surveys continue to project enormous demand, often projecting that at least half of the households, or half of the people, want wireless personal communications. Trade magazines, newsletters, conferences, and seminars on the subject by many different names have become too numerous to keep track of, and technical journals, magazines, conferences and symposia continue to proliferate and to have ever increasing attendance and numbers of papers presented. It is clear that wireless personal communications is, by any measure, the fastest growing segment of telecommunications.

However, if you look carefully at the seemingly endless discussions of the topic, you cannot help but note that they are often describing different "things", i.e., different versions of wireless personal communications [1, 2]. Some discuss pagers, or messaging, or data systems, or access to the National Information Infrastructure, while others emphasize cellular radio, or cordless telephones, or dense systems of satellites. Many make reference to popular fiction entities like Dick Tracy, Maxwell Smart, or *Star Trek*.

Thus, it appears that almost everyone wants Wireless Personal Communications, but, *What Is It*?!! There are many different ways to segment the complex topic into different communications applications, modes, functions, extent of coverage, or mobility [1, 2]. The complexity of the issues has resulted in considerable confusion in the industry, as evidenced by the many different wireless systems, technologies, and services being offered, planned, or proposed. Many different industry groups and regulatory entities are becoming involved. The confusion is a natural consequence of the massive dislocations that are occurring, and will continue to occur, as we progress along this large change in the paradigm of the way we communicate. Among the different changes that are occurring in our communications paradigm, perhaps the major ingredient is the change from wired fixed place-to-place communications to wireless mobile person-to-person communications. Within this major change are also many other changes, e.g., an increase in the significance of data and message communications, a perception of possible changes in video applications, and changes in the regulatory and political climates.

This article attempts to identify different issues and to put many of the activities in wireless into a framework that can provide perspective on what is driving them, and perhaps even yield some indication of where they appear to be going in the future. However, like any attempt to categorize many complex interrelated issues, there are some that don't quite fit into neat categories, so there will remain some dangling loose ends. Like any major paradigm shift, there will continue to be considerable confusion as many entities attempt to interpret the different needs and expectations associated with the new paradigm.

Background and Issues

Mobility and Freedom from Tethers

Perhaps the clearest ingredients in all of the wireless personal communications activity are the desire for mobility in communications, and the companion desire to be free from tethers, i.e., from physical connections to communications networks. These desires are clear from the very rapid growth of mobile technologies that provide primarily two-way voice services, even though economical wireline voice services are readily available. For example, cellular mobile radio has experienced rapid growth. Growth rates have been between 35 and 60 percent per year in the United States for a decade, with the total number of subscribers reaching 20 million by year-end 1994. The often neglected wireless companions to cellular radio, i.e., cordless telephones, have experienced even more rapid, but harder to quantify, growth with sales rates often exceeding 10 million sets a year in the United States, and with an estimated usage significantly exceeding 50 million in 1994. Telephones in airliners, have also become commonplace. Similar, or even greater, growth in these wireless technologies has been experienced throughout the world.

Paging and associated messaging, while not providing two-way voice, do provide a form of tetherless mobile communications to many subscribers worldwide. These services have also experienced significant growth. There is even a glimmer of a market in the many different specialized wireless data applications evident in the many wireless local area network (WLAN) products on the market, the several wide area data services being offered, and the specialized satellite-based message services being provided to trucks on highways.

1070-9916/95/$04.00 © 1995 IEEE

The topics discussed in the previous two paragraphs indicate a dominant issue separating the different evolutions of wireless personal communications. That issue is the voice versus data communications issue that permeates all of communications today; this division also is very evident in fixed networks. The packet-oriented computer communications community and the circuit-oriented voice telecommunications (telephone) community hardly talk to each other, and often speak different languages in addressing similar issues. Although they often converge to similar overall solutions at large scales (e.g., hierarchical routing with exceptions for embedded high usage routes), the small scale initial solutions are frequently quite different. Asynchronous Transfer Mode (ATM)-based networks are an attempt to integrate, at least partially, the needs of both the packet-data and circuit-oriented communities.

Superimposed on the voice-data issue is an issue of competing modes of communications that exist in both fixed and mobile forms. These different modes include:

Messaging, where the communication is not real time, but is by way of message transmission, storage, and retrieval. This mode is represented by voice mail, electronic facsimile (fax), and electronic mail (e-mail), the latter of which appears to be a modern automated version of an evolution that includes telegraph and telex. Radio paging systems often provide limited one-way messaging, ranging from transmitting only the number of a calling party, to longer alpha-numeric text messages.

Real-time two-way communications, represented by the telephone, cellular mobile radio telephone, and interactive text (and graphics) exchange over data networks. Two-way video phone always captures significant attention and fits into this mode; however, its benefit/cost ratio has yet to exceed a value that customers are willing to pay.

Paging, i.e., broadcast with no return channel, alerts a paged party that someone wants to communicate with him/her. Paging is like the ringer on a telephone, without having the capability for completing the communications.

Agents, new high level software applications or entities being incorporated into some computer networks. When launched into a data network, an "agent" is aimed at finding information by some title or characteristic, and returning the information to the point from which the agent was launched.

There are still other ways in which wireless communications have been segmented in attempts to optimize a technology to satisfy the needs of some particular group. Examples include:

- User location, that can be differentiated by indoors or outdoors, or on an airplane or a train.
- Degree of mobility, that can be differentiated either by speed, e.g., vehicular, pedestrian, or stationary, or by size of area throughout which communications are provided.

At this point one should again ask: "Wireless Personal Communications — *What Is It*?!!" The evidence suggests that what is being sought by users, and produced by providers, can be categorized according to the following two main characteristics.

Communications Portability and Mobility on many different scales:

- Within a house or building (cordless telephone, wireless local area networks (WLANs)).
- Within a campus, a town, or a city (cellular radio, WLANs, wide area wireless data, radio paging, extended cordless telephone).
- Throughout a state or region (cellular radio, wide area wireless data, radio paging, satellite-based wireless).
- Throughout a large country or continent (cellular radio, paging, satellite-based wireless).
- Throughout the world?!!

Communications by many different modes for many different applications:

- Two-way voice.
- Data.
- Messaging.
- Video?

Thus, it is clear why wireless personal communications today is not one technology, not one system, and not one service, but encompasses many technologies, systems and services optimized for different applications.

Evolution of Technologies, Systems, and Services

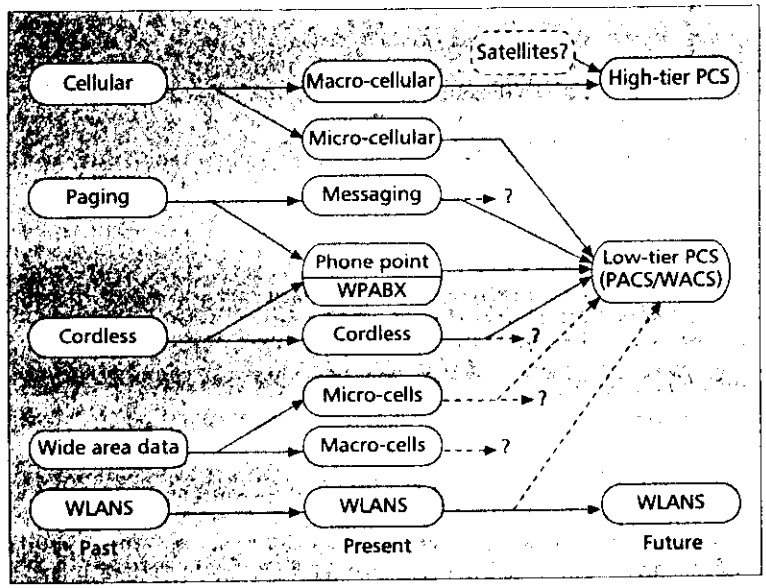
Technologies and systems [1-7] that are currently providing, or are proposed to provide, wireless communications services can be grouped into about seven relatively distinct groups, although there may be some disagreement on the group definitions, and in what group some particular technology or system belongs. All of the technologies and systems are evolving as technology advances and perceived needs change. Some trends are becoming evident in the evolutions. In this section, different groups and evolutionary trends are explored along with factors that influence the characteristics of members of the groups. The grouping is generally with respect to scale of mobility and communications applications or modes.

Perhaps the clearest ingredients in all of the wireless personal communications activity are the desire for mobility in communications, and the companion desire to be free from tethers, i.e., from physical connections to communications networks.

Cordless Telephones

Cordless telephones [1-3] generally can be categorized as providing low mobility, low-power, two-way tetherless voice communications, with low mobility applying both to the range and the user's speed. Cordless telephones using analog radio technologies appeared in the late 1970s, and have experienced spectacular growth. They have evolved to digital radio technologies in the forms of second-generation cordless telephone (CT-2), and Digital European Cordless Telephone (DECT) standards in Europe, and several different Industrial Scientific Medical (ISM) band technologies in the United States.[1]

[1] *These ISM technologies either use spread spectrum techniques (direct sequence or frequency hopping), or very low transmitter power (≤ 1 mw) as required by the ISM band regulations.*



■ **Figure 1.** *Digital wireless access systems evolution.*

Cordless telephones were originally aimed at providing economical, tetherless voice communications inside residences, i.e., at using a short wireless link to replace the cord between a telephone base unit and its handset. The most significant considerations in design compromises made for these technologies are to minimize total cost, while maximizing the "talk time" away from the battery charger. For digital cordless phones intended to be carried away from home in a pocket, e.g., CT-2 or DECT, handset weight and size are also major factors. These considerations drive designs toward minimizing complexity, and minimizing the power used for signal processing and for transmitting.

Cordless telephones compete with wireline telephones. Therefore, high circuit quality has become a requirement. Early cordless sets had marginal quality. They were purchased by the millions, and discarded by the millions, until manufacturers produced higher-quality sets. Cordless telephones sales then exploded. Their usage has become commonplace, approaching, and perhaps exceeding, usage of "corded" telephones.

The compromises accepted in cordless telephone design in order to meet the cost, weight, and talk-time objectives are:

- Few users per MHz.
- Few users per base unit (many link together a particular handset and base unit).
- Large number of base units per unit area; one or more base units per wireline access line (in high-rise apartment buildings the density of base units is very large).
- Short transmission range.

There is no added network complexity since a base unit looks to a telephone network like a wireline telephone. These issues are also discussed in [1, 2].

Digital cordless telephones in Europe have been evolving for a few years to extend their domain of use beyond the limits of inside residences. Cordless telephone, second generation, (CT-2) has evolved to provide telepoint or phone-point services. Base units are located in places where people congregate, e.g., along city streets and in shopping malls, train stations, etc. Handsets registered with the phone-point provider can place calls when within range of a telepoint. CT-2 does not provide capability for transferring (handing off) active wireless calls from one phone point to another if a user moves out of range of the one to which the call was initiated. A CT-2+ technology, evolved from CT-2 and providing limited handoff capability, is being deployed in Canada. Phone-point service was introduced in the United Kingdom twice, but failed to attract enough customers to become a viable service. However, in Singapore and Hong Kong, CT-2 phone-point has grown rapidly, reaching over 150,000 subscribers in Hong Kong [8] in mid-1994. The reasons for the success in some places and failure in others are still being debated, but it is clear that the compactness of the Hong Kong and Singapore populations make the service more widely available, using fewer base stations than in more spreadout cities. Complaints of CT-2 phone-point users in trials have been that the radio coverage was not complete enough, and/or they could not tell whether there was coverage at a particular place, and the lack of handoff was inconvenient. In order to provide the "alerting" or "ringing" function for phone-point service, conventional radio pagers have been built into some CT-2 handsets. (The telephone network to which a CT-2 phone point is attached has no way of knowing from which base units to send a ringing message, even though the CT-2 handsets can be "rung" from a home base unit).

Another European evolution of cordless telephones is Digital European Cordless Telephone (DECT) which was optimized for use inside buildings. Base units are attached through a controller to private branch exchanges (PBXs), key telephone systems, or phone company CENTREX telephone lines. DECT controllers can hand off active calls from one base unit to another as users move, and can "page" or "ring" handsets as a user walks through areas covered by different base units.

These cordless telephone evolutions to more widespread usage outside and inside with telepoints, and to usage inside large buildings are illustrated in Fig. 1, along with the integration of paging into handsets to provide alerting for phone-point services. They represent the first attempts to increase the service area of mobility for low-power cordless telephones.

Some of the characteristics of the digital cordless telephone technologies, CT-2 and DECT, are listed in Table 1. Additional information can be found in References [2, 3]. Even though there are significant differences between these technologies, e.g., multiple access technology (FDMA or TDMA/FDMA), and channel bit rate, there are many similarities that are fundamental to the design objectives discussed earlier, and to a user's perception of them. These similarities and their implications are as follows.

32 kb/s adaptive differential pulse code modulation (ADPCM) digital speech encoding: this is a low complexity (low signal processing power) speech encoding process that provides wireline speech quality and is an international standard.

Average transmitter power ≤ 10 milliwatts: this permits many hours of talk time with small, low-cost, lightweight batteries, but provides limited radio range.

	High Power Systems				Low Power Systems			
	Digital Cellular (High Tier PCS)				Low Tier PCS		Digital Cordless	
System	IS-54	IS-95 (DS)	GSM	DCS-1800	WACS/PACS	Handi-Phone	DECT	CT-2
Multiple access	TDMA/FDMA	CDMA/FDMA	TDMA/FDMA	TDMA/FDMA	TDMA/FDMA	TDMA/FDMA	TDMA/FDMA	FDMA
Freq. band (MHz) Uplink (MHz) Downlink (MHz)	869-894 824-849 (USA)	869-894 824-849 (USA)	935-960 890-915 (Eur.)	1710-1785 1805-1880 (UK)	Emerg. Tech.* (USA)	1895- 1907 (Japan)	1880- 1900 (Eur.)	864-868 (Eur. and Asia)
RF ch. spacing Downlink (KHz) Uplink (KHz)	30 30	1250 1250	200 200	200 200	300 300	300	1728	100
Modulation	π/4 DQPSK	BPSK/QPSK	GMSK	GMSK	π/4 QPSK	π/4 DQPSK	GFSK	GFSK
Portable txmit Power, max./avg.	600 mW/ 200 mW	600 mW	1 W/ 125 mW	1 W/ 125 mW	200 mW/ 25 mW	80 mW/ 10 mW	250 mW/ 10 mW	10 mW/ 5 mW
Speech coding	VSELP	QCELP	RPE-LTP	RPE-LTP	ADPCM	ADPCM	ADPCM	ADPCM
Speech rate (kb/s)	7.95	8 (var.)	13	13	32/16/8	32	32	32
Speech ch./RF ch.	3	-	8	8	8/16/32	4	12	
Ch. bit rate (kb/s) Uplink (kb/s) Downlink (kb/s)	48.6 48.6		270.833 270.833	270.833 270.833	384 384	384	1152	72
Ch. coding	1/2 rate conv.	1/2 rate fwd 1/3 rate rev.	1/2 rate conv.	1/2 rate conv.	CRC	CRC	CRC (control)	
Frame (ms)	40	20	4.615	4.615	2.5	5	10	2

* Spectrum is 1.85 to 2.2 GHz allocated by the FCC for emerging technologies; DS is direct sequence.

■ **Table 1.** *Wireless PCS technologies.*

Low-complexity radio signal processing: there is no forward error correction and no complex multipath mitigation (i.e., no equalization or spread spectrum).

Low transmission delay, e.g., < 50 ms, and for CT-2 < 10 ms round trip: this is a speech-quality and network-complexity issue. A maximum of 10 ms should be allowed, taking into account additional inevitable delay in long-distance networks. Echo cancellation is generally required for delays > 10 ms.

Simple frequency-shift modulation and noncoherent detection: while still being low in complexity, the slightly more complex 4QAM modulation with coherent detection provides significantly more spectrum efficiency, range and interference immunity.

Dynamic channel allocation: While this technique has potential for improved system capacity, the cordless-telephone implementations do not take full advantage of this feature for handoff, and thus cannot reap the full benefit for moving users [9, 10].

Time division duplex (TDD): this technique permits the use of a single contiguous frequency band, and implementation of diversity from one end of a radio link. However, unless all base station transmissions are synchronized in time, it can incur severe cochannel interference penalties in outside environments [9, 11]. Of course, for cordless telephones used inside with base stations not having a propagation advantage, this is not a problem. Also, for small indoor PBX networks, synchronization of base station transmission is easier than is synchronization throughout a widespread outdoor network, which can have many adjacent base stations connected to different geographic locations for central control and switching.

Cellular Mobile Radio Systems

Cellular mobile radio systems are becoming known in the United States as high-tier Personal Communications Service (PCS), particularly when implemented in the new 1.9 GHz PCS bands [12]. These systems generally can be categorized as providing high-mobility, wide-ranging, two-way tetherless voice communications. In these systems, high mobility refers to vehicular speeds, and also to widespread regional to nationwide coverage [1, 2, 7]. Mobile radio has been evolving for over 50 years. Cellular radio integrates wireless access with large-scale networks having sophisticated intelligence to manage mobility of users.

Cellular radio was designed to provide voice service to wide-ranging vehicles on streets and highways [1-3, 13], and generally uses transmitter power on the order of 100 times that of cordless telephones (≈ 2 watts for cellular). Thus, cellular systems can only provide reduced service to handheld sets that are disadvantaged by using somewhat lower transmitter power (< 0.5 watts) and less efficient antennas than vehicular sets. Handheld sets used inside buildings have the further disadvantage of attenuation through walls that is not taken into account in system design.

Cellular radio or high-tier PCS has experienced large growth as noted earlier. In spite of the limitations on usage of handheld sets noted above, handheld cellular sets have become very popular,

with their sales becoming comparable to the sales of vehicular sets. Frequent complaints from handheld cellular users are that batteries are too large and heavy, and both talk time and standby time are inadequate.

Cellular radio at 800 MHz has evolved to digital radio technologies [1-3] in the forms of the deployed systems standards:

- Global Standard for Mobile (GSM) in Europe.
- Japanese or Personal Digital Cellular (JDC or PDC) in Japan.
- U.S. TDMA digital cellular known as USDC or IS-54.

and in the form of the code division multiple access (CDMA) standard, IS-95, which is under development, but not yet deployed.

The most significant consideration in the design compromises made for the U.S. digital cellular or high-tier PCS systems was the high cost of cell sites (base stations). A figure often quoted is U.S. $1 million for a cell site. This consideration drove digital system designs to:

- Maximize users per MHz.
- Maximize the users per cell site.

Because of the need to cover highways running through low population-density regions between cities, the relatively high transmitter power requirement was retained to provide maximum range from high antenna locations.

Compromises that were accepted while maximizing the above parameters are:

- High transmitter power consumption.
- High user-set complexity, and thus high signal-processing power consumption.
- Low circuit quality.

The use of microcell base stations provides large increases in overall system capacity, while also reducing the cost per available radio channel, and the battery drain on portable subscriber equipment.

- High network complexity, e.g., the new IS-95 technology will require complex new switching and control equipment in the network, as well as high-complexity wireless-access technology.

Cellular radio or high-tier PCS has also been evolving for a few years in a different direction, toward very small coverage areas or microcells. This evolution provides increased capacity in areas having high user density, as well as improved coverage of shadowed areas. Some microcell base stations are being installed inside, in conference center lobbies and similar places of high user concentrations. Of course, microcells also permit lower transmitter power that conserves battery power when power control is implemented, and base stations inside buildings circumvent the outside wall attenuation. Low complexity microcell base stations also are considerably less expensive than conventional cell sites, perhaps two orders of magnitude less expensive. Thus, the use of microcell base stations provides large increases in overall system capacity, while also reducing the cost per available radio channel, and the battery drain on portable subscriber equipment. This microcell evolution, illustrated in Fig. 1, moves handheld cellular sets in a direction similar to that of the expanded-coverage evolution of cordless telephones to phone points and wireless PBX.

Some of the characteristics of digital-cellular or high-tier PCS technologies are listed in Table 1 for IS-54, IS-95, and GSM at 900 MHz, and DCS-1800, which is GSM at 1800 MHz. Additional information can be found in [1-3]. The JDC or PDC technology, not listed, is similar to IS-54. As with the digital cordless technologies, there are significant differences among these cellular technologies, e.g., modulation type, multiple access technology, and channel bit rate. However, there are also many similarities that are fundamental to the design objectives discussed earlier. These similarities and their implications are as follows.

Low bit-rate speech coding; ≤ 13 kb/s with some ≤ 8 kb/s: low bit-rate speech coding obviously increases the number of users per MHz and per cell site. However, it also significantly reduces speech quality [1], and does not permit the tandemming of speech encoding while traversing a network. That is, when low bit rate speech is transcoded to a different encoding format, e.g., to 64 kb/s as is used in many networks, or from an IS-54 phone on one end to a GSM or IS-95 phone on the other end, the speech quality deteriorates precipitously. While this may not be a serious issue for a vehicular mobile user who has no choice other than not to communicate at all, it is likely to be a serious issue in an environment where a wireline telephone is available as an alternative. It is also less serious when there are few mobile-to-mobile calls through the network, but, as wireless usage increases, and digital mobile-to-mobile calls become commonplace, the marginal transcoded speech quality is likely to become a serious issue.

Some implementations make use of speech inactivity: this further increases the number of users per cell site, i.e., the cell-site capacity. However, it also further reduces speech quality [1] because of the difficulty of detecting the onset of speech. This problem is even worse in an acoustically noisy environment like an automobile.

High transmission delay; ≈ 200 ms round trip: this is another important circuit-quality issue. Such large delay is about the same as one-way transmission through a synchronous-orbit communications satellite. A voice circuit with digital cellular technology on both ends will experience the delay of a full satellite circuit. It should be recalled that one reason long-distance circuits have been removed from satellites and put onto fiber-optic cable is because customers find the delay to be objectionable. This delay in digital cellular technology results from both computation for speech bit-rate reduction, and from complex signal processing, e.g., bit interleaving, error correction decoding, and multipath mitigation (equalization or spread spectrum (CDMA)).

High-complexity signal processing, both for speech encoding and for demodulation: signal processing has been allowed to grow without bound, and is about a factor of 10 greater than that used in the low-complexity digital cordless telephones [1]. Since several watts are required from a battery to produce the high transmitter power in a cellular or high-tier PCS set, signal-processing power is not as significant as it is in the low-power cordless telephones.

82

Fixed channel allocation: the difficulties associated with implementing capacity-increasing dynamic channel allocation to work with handoff [9, 10] have impeded its adoption in systems requiring reliable and frequent handoff.

Frequency division duplex (FDD): cellular systems have already been allocated paired-frequency bands suitable for FDD. Thus, the network or system complexity required for providing synchronized transmissions [9, 11] from all cell sites for TDD has not been embraced in these digital cellular systems. Note that TDD has not been employed in IS-95 even though such synchronization is required for other reasons.

Mobile/portable set power control: the benefits of increased capacity from lower overall co-channel interference, and reduced battery drain have been sought by incorporating power control in the digital cellular technologies.

Wide Area Wireless Data Systems

Existing wide area data systems generally can be categorized as providing high mobility, wide-ranging, low-data-rate digital data communications to both vehicles and pedestrians [1, 2]. These systems have not experienced the rapid growth that the two-way voice technologies have, even though they have been deployed in many cities for a few years and have established a base of customers in several countries. Examples of these packet data systems are shown in Table 2.

The earliest and best known of these systems in the United States are the ARDIS network developed and run by Motorola, and the RAM mobile data network based on Ericsson Mobitex Technology. These technologies were designed to make use of standard, two-way voice, land mobile-radio channels, with 12.5 KHz or 25 kHz channel spacing. In the United States these are specialized mobile radio services (SMRS) allocations around 450 MHz and 900 MHz. Initially, the data rates were low: 4.8 kb/s for ARDIS and 8 kb/s for RAM. The systems use high transmitter power (several tens of watts) to cover large regions from a few base stations having high antennas. The relatively low data capacity of a relatively expensive base station has resulted in economics that have not favored rapid growth.

The wide area mobile data systems also are evolving in several different directions in an attempt to improve base station capacity, economics, and the attractiveness of the service. The technologies used in both the ARDIS and RAM networks are evolving to higher channel bit rates of 19.2 kb/s.

The cellular carriers and several manufacturers in the United States are developing and deploying a new wide area packet data network as an overlay to the cellular radio networks. This Cellular Digital Packet Data (CDPD) technology shares the 30 kHz spaced 800 MHz voice channels used by the analog FM Advanced Mobile Phone Service (AMPS) systems. Data rate is 19.2 kb/s. The CDPD base station equipment also shares cell sites with the voice cellular radio system. The aim is to reduce the cost of providing packet data service by sharing the costs of base stations with the better-established and higher cell-site capacity cellular systems. This is a strategy similar to that used by nationwide fixed wireline packet-data-networks that could not provide an economically viable data service if they did not share costs by leasing a small amount of the capacity of the interexchange networks that are paid for largely by voice traffic.

	CDPD	RAM Mobile (Mobitex)	ARDIS (KDT)	Metricom (MDN)
Data rate	19.2 KB/s	8 Kb/s [19.2 Kb/s]	4.8 Kb/s [19.2 Kb/s]	~ 76 Kb/s
Modulation	GMSK BT = 0.5	GMSK	GMSK	GMSK
Frequency	~ 800 MHz	~ 900 MHz	~ 800 MHz	~ 915 MHz
Chan. spacing	30 KHz	12.5 KHz	25 KHz	160 KHz
Status	1994 service	Full service	Full service	In service
Access means	Unused AMPS channels	Slotted Aloha CSMA		FH SS (ISM)
Transmit power			40 watt	1 watt

Note: data in square brackets [ ] indicates proposed.
CDPD: Cellular Digital Packet Data
MDN: Microcellular Data Network
ARDIS: Advanced Radio Data Information Service

■ **Table 2.** *Wide area wireless packet data systems.*

Another evolutionary path in wide area wireless packet data networks is toward smaller coverage areas or microcells. This evolutionary path also is indicated on Fig. 1. The microcell data networks are aimed at stationary or low-speed users. The design compromises are aimed at reducing service costs by making very small and inexpensive base stations that can be attached to utility poles, the sides of buildings, and inside buildings, and can be widely distributed throughout a region. Base-station-to-base-station wireless links are used to reduce the cost of the interconnecting data network. In one network this decreases the overall capacity to serve users, since it uses the same radio channels that are used to provide service. Capacity is expected to be made up by increasing the number of base stations that have connections to a fixed-distribution network as service demand increases. Another such network uses other dedicated radio channels to interconnect base stations. In the high-capacity limit, these networks will look more like a conventional cellular network architecture, with closely spaced, small, inexpensive base stations, i.e., microcells, connected to a fixed infrastructure. Specialized wireless data networks have been built to provide metering and control of electric power distribution, e.g., Celldata, and Metricom in California.

A large microcell network of small inexpensive base stations has been installed in the lower San Francisco Bay Area by Metricom, and public packet-data service was offered during early 1994. Most of the small (shoe-box-size) base stations are mounted on street light poles. Reliable data rates are about 75 kb/s. The technology is based on slow frequency-hopped spread spectrum in the 902-928 MHz U.S. Industrial Scientific Medical (ISM) band. Transmitter power is 1 watt maximum, and power control is used to minimize interference and maximize battery life time.

High-Speed Wireless Local-Area Networks (WLANs)

Wireless local-area data networks (WLANs) can be categorized as providing low-mobility high-speed data communications within a confined region, e.g., a campus or a large building. Coverage range from a wireless data terminal is short, tens to hundreds of feet, like cordless telephones. Coverage is limited to within a room or to several rooms in a building. WLANs have been evolving for a few years, but overall, the situation is chaotic, with many different products being offered by many different vendors [1, 6]. There is no stable definition of the needs or design objectives for WLANs, with data rates ranging from hundreds of kb/s to more than 10 MB/s, and with several products providing one or two MB/s wireless link rates. The best description of the WLAN evolutionary process is: "having severe birth pains." An IEEE standards committee, 802.11, has been attempting to put some order into this topic, but their success has been

Product Company Location	Freq. (MHz)	Link rate	User rate	Protocol(s)	Access	No. of chan. or spread factor	Mod./coding	Power	Network topology
Altair Plus II Motorola Arlington Hts., IL	18-19 GHz	15 Mb/s	5.7 Mb/s	Ethernet			4-level FSK	25 mW peak	Eight devices/ radio; radio to base to Ethernet
WaveLAN NCR/AT&T Dayton, OH	902-928	2 Mb/s	1.6 Mb/s	Ethernet-like	DS SS		DQPSK	250 mW	Peer-to-peer
AirLAN Solectek San Diego, CA	902-928		2 Mb/s	Ethernet	DS SS		DQPSK	250 mW	PCMCIA w/ant; radio to hub
Freeport Windata Inc. Northboro, MA	902-928	16 Mb/s	5.7 Mb/s	Ethernet	DS SS	32 chips/bit	16 PSK trellis coding	650 mW	Hub
Intersect Persoft Inc. Madison, WI	902-928		2 Mb/s	Ethernet, token-ring	DS SS		DQPSK	250 mW	Hub
LAWN O'Neill Comm. Horsham, PA	902-928		38.4 kb/s	AX.25	SS	20 users/chan.; max. 4 chan.		20 mW	Peer-to-peer
WILAN WILAN Inc. Calgary, Alberta	902-928	20 Mb/s	1.5 Mb/s/ chan.	Ethernet, token ring	CDMA/ TDMA	3 chan/ 10-15 links each	"unconventional"	30 mW	Peer-to-peer
RadioPort ALPS Electric USA	902-928		242 kb/s	Ethernet	SS	?/3 channels		100 mW	Peer-to-peer
ArLAN 600 Telesys. SLW Don Mills, Ont.	902-928; 2.4 GHz		1.35 Mb/s	Ethernet	SS			1 W max	PCs with ant.; radio to hub
RadioLink Cal. Microwave Sunnyvale, CA	902-928; 2.4 GHz	250 kb/s	64 kb/s		FH SS	250 ms/hop 500 kHz space			Hub
Range LAN Proxim, Inc. Mountain View, CA	902-928		242 kb/s	Ethernet, token ring	DS SS	3 chan.		100 mW	
RangeLAN2 Proxim, Inc. Mountainview, CA	2.4 GHz	1.6 Mb/s	50 kb/s max.	Ethernet, token ring	FH SS	10 chan. @ 5 kb/s; 15 sub-ch. each		100 mW	Peer-to-peer bridge
Netwave Xircom Calabasas, CA	2.4 GHz	1 Mb/s/ adaptor		Ethernet, token ring	FH SS	82 1-MHz chan. or "hops"			Hub
Freelink Cabletron Sys. Rochester, NH	2.4 and 5.8 GHz		5.7 Mb/s	Ethernet	DS SS	32 chips/bit	16 PSK trellis coding	100 mW	Hub

■ **Table 3.** *Partial list of WLAn products.*

somewhat limited. A partial list of some advertised products is given in Table 3. Users of WLANs are not nearly as numerous as the users of more voice-oriented wireless systems. Part of the difficulty stems from these systems being driven by the computer industry that views the wireless system as just another plug-in interface card, without giving sufficient consideration to the vagaries and needs of a reliable radio system.

There are two overall network architectures pursued by WLAN designers. One is a centrally coordinated and controlled network that resembles other wireless systems. There are base stations in these networks that exercise overall control over channel access [14].

The other type of network architecture is the self organizing and distributed controlled network where every terminal has the same function as every other terminal, and networks are formed ad-hoc by communications exchanges among terminals. Such ad-hoc networks are more like citizen band (CB) radio networks, with similar expected limitations if they were ever to become very widespread. Nearly all WLANs in the United States have attempted to use one of the ISM frequency bands for unlicensed operation under part 15 of the FCC rules. These bands are 902 to 928 MHz, 2400 to 2483.5 MHz, and 5725 to 5850 MHz, and they require users to accept interference from any interfering source that may also be using the frequency. The use of ISM bands has further handicapped WLAN development because of the requirement for use of either frequency hopping or direct sequence spread spectrum as an access technology, if transmitter power is to be adequate to cover more than a few feet. One exception to the ISM band implementations is the Motorola ALTAIR, which operates in a licensed band at 18 GHz. The technical and economic challenges of operation at 18 GHz have hampered the adoption of this 10 to 15 MB/s technology. The frequency-spectrum constraints have been improved in the United States with the recent FCC allocation of spectrum from 1910 to 1930 MHz for unlicensed "data PCS" applications. Use of this new spectrum requires implementation of an access "etiquette" incorporating "Listen before Transmit" in an attempt to provide some coordination of an otherwise potentially chaotic, uncontrolled environment [15]. Also, since spread spectrum is not a requirement, access technologies and multipath mitigation techniques more compatible with the needs of packet data transmission [6], e.g., multipath equalization or multicarrier transmission can be incorporated into new WLAN designs.

Three other widely different WLAN activities also need mentioning. One is a large European Telecommunications Standards Institute(ETSI) activity to produce a standard for High Performance Radio Local Area Network (HIPERLAN), a 20 MB/s WLAN technology to operate near 5 GHz. Other activities are large, U.S. Advance Research Projects Agency (ARPA)-sponsored, WLAN research projects at the Universities of California at Berkeley (UCB), and at Los Angeles (UCLA). The UCB Infopad project is based on a coordinated network architecture with fixed coordinating nodes and direct-sequence spread spectrum (CDMA), whereas, the UCLA project is aimed at peer-to-peer networks and uses frequency hopping. Both ARPA sponsored projects are concentrated on the 900 MHz ISM band.

As computers shrink in size from desktop, to laptop, to palmtop, mobility in data network access is becoming more important to the user. This fact, coupled with the availability of more usable frequency spectrum, and perhaps some progress on standards, may speed the evolution and adoption of wireless mobile access to WLANs. From the large number of companies making products, it is obvious that many believe in the future of this market.

As computers shrink in size from desktop, to laptop, to palmtop, mobility in data network access is becoming more important to the user. This fact, coupled with the availability of more usable frequency spectrum, and perhaps some progress on standards, may speed the evolution and adoption of wireless mobile access to WLANs.

Paging/Messaging Systems

Radio paging began many years ago as a "one bit" messaging system. The one bit was "some one wants to communicate with you." More generally, paging can be categorized as one-way messaging over wide areas. The one-way radio link is optimized to take advantage of the asymmetry. High transmitter power (hundreds of watts to kilowatts), and high antennas at the fixed base stations permit low complexity, very-low-power-consumption, pocket paging receivers that provide long usage time from small batteries. This combination provides the large radio-link margins needed to penetrate walls of buildings without burdening the user set battery. Paging has experienced steady rapid growth for many years and serves about 15 million subscribers in the United States

Paging also has evolved in several different directions. It has changed from analog tone coding for user identification to digitally encoded messages. It has evolved from the one-bit message, "someone wants you," to multibit messages from, first, the calling party's telephone number to, now, short e-mail text messages. This evolution is noted in Fig. 1.

The region over which a page is transmitted has also increased from a) local, around one transmitting antenna; to b) regional, from multiple widely-dispersed antennas; to c) nationwide, from large networks of interconnected paging transmitters. The integration of paging with CT-2 user sets for phone-point call alerting was noted previously.

Another "evolutionary" paging route sometimes proposed is "two-way" paging. However, this is an ambiguous and unrealizable concept, since the requirement for two-way communications destroys the asymmetrical link advantage so well exploited by paging. "Two-way" paging puts a transmitter in the user's set, and brings along with it all the design compromises that must be faced in such a two-way radio system. Thus, the word "paging" is not appropriate to describe a system that provides two-way communications.

Satellite-Based Mobile Systems

Satellite-based systems are the epitome of wide-area-coverage, expensive, base station systems. They generally can be categorized as providing two-way (or one-way) limited quality voice, and/or very limited data or messaging, to very wide-ranging vehicles (or fixed locations). These systems can provide very widespread, often global, coverage, e.g., to ships at sea by INMARSAT. There are a few messaging systems in operation, e.g., to trucks on highways in the United States by Qualcomm's Omnitracs system.

It remains to be seen whether there will be enough users with enough money in low population density regions of the world to make satellite mobile systems economically viable.

A few large scale mobile satellite systems have been proposed and are being pursued; perhaps the best known is Motorola's Iridium, and others include Odyssey, Globalstar, and Teledesic. The strength of satellite systems is their ability to provide large regional or global coverage to users outside buildings. However, it is very difficult to provide adequate link margin to cover inside buildings, or even to cover locations shadowed by buildings, trees or mountains. A satellite system's weakness is also its large coverage area. It is very difficult to provide from earth orbit the small coverage cells that are necessary for providing high overall systems capacity from frequency reuse. This fact, coupled with the high cost of the orbital base stations, results in low capacity along with the wide overall coverage, but also in expensive service. Thus, satellite systems are not likely to compete favorably with terrestrial systems in populated areas, or even along well traveled highways. They can complement terrestrial cellular or PCS systems in low population density areas. It remains to be seen whether there will be enough users with enough money in low population density regions of the world to make satellite mobile systems economically viable.

Proposed satellite systems range from a) low-earth-orbit (LEOS) systems, having tens to hundreds of satellites, through b) intermediate or medium height systems (MEOS?), to c) geostationary or geosynchronous orbit systems (GEOS), having fewer than ten satellites. LEOS require more, but less expensive, satellites to cover the earth, but they can more easily produce smaller coverage areas, and thus provide higher capacity within a given spectrum allocation. Also, their transmission delay is significantly less (perhaps two orders of magnitude!), providing higher-quality voice links as discussed previously. On the other hand, GEOs require only a few, somewhat more expensive, satellites (perhaps only three), and are likely to provide lower capacity within a given spectrum allocation, and suffer severe transmission-delay impairment on the order of 0.5 seconds. Of course, MEOS fall in-between these extremes. The possible evolution of satellite systems to complement high tier PCS is indicated in Fig. 1.

Evolution Toward the Future and To Low-Tier Personal Communications Services

After looking at the evolution of several wireless technologies and systems in the previous sections, it appears appropriate to ask again: "Wireless Personal Communications — What Is It?" All of the technologies in the previous sections claim to provide wireless personal communications, and all do to some extent. However, all have significant limitations and all are evolving in attempts to overcome the limitations. It seems appropriate to ask, what are the likely endpoints? Perhaps some hint of the endpoints can be found by exploring what users see as limitations of existing technologies and systems, and by looking at the evolutionary trends.

In order to do so, we summarize some important clues from the previous sections, and project them, along with some U.S. standards activity, toward the future.

Digital Cordless Telephones

- Strengths: good circuit quality; long talk time; small lightweight battery; low-cost sets and service.
- Limitations: limited range; limited usage regions.
- Evolutionary trends: phone-points in public places; wireless PBX in business.
- Remaining limitations and issues: limited usage regions and coverage holes; limited or no handoff; limited range.

Digital Cellular Pocket Handsets

- Strength: widespread service availability.
- Limitations: limited talk time; large heavy batteries; high-cost sets and service; marginal circuit quality; holes in coverage and poor in-building coverage; limited data capabilities; complex technologies.
- Evolutionary trends: microcells to increase capacity and in building coverage, and to reduce battery drain; satellite systems to extend coverage.
- Remaining limitations and issues: limited talk time and large battery; marginal circuit quality; complex technologies.

Wide Area Data

- Strength: digital messages.
- Limitations: no voice; limited data rate; high cost.
- Evolutionary trends: microcells to increase capacity and reduce cost; share facilities with voice systems to reduce cost.
- Remaining limitations and issues: no voice; limited capacity.

Wireless Local Area Networks (WLANs)

- Strength: high data rate.
- Limitations: insufficient capacity for voice; limited coverage; no standards; chaos.
- Evolutionary trends: hard to discern from all the churning.

Paging/messaging

- Strengths: widespread coverage; long battery life; small lightweight sets and batteries; economical.
- Limitations: one-way message only; limited capacity.
- Evolutionary desire: two-way messaging and/or voice; capacity.
- Limitations and issues: two-way link cannot exploit the advantages of one-way link asymmetry.

There is a strong trajectory evident in these systems and technologies, aimed at providing the following features.

High Quality Voice and Data

- To small, lightweight, pocket carried communicators.
- Having small lightweight batteries.
- Having long talk time, and long standby battery life.
- Providing service over large coverage regions.
- For pedestrians in populated areas (but not requiring high population density).
- Including low to moderate speed mobility with handoff.

Economical Service

- Low subscriber-set cost.
- Low network-service cost.

Privacy and Security of Communications

- Encrypted radio links.

This trajectory is evident in all of the evolving technologies, but can only be partially satisfied by any of the existing and evolving systems and technologies! Trajectories from all of the evolving technologies and systems are illustrated in Fig. 1 as being aimed at low-tier personal communications systems or services, i.e., low-tier PCS.

Taking characteristics from cordless, cellular, wide area data and, at least moderate-rate, WLANs, suggests the following attributes for this low-tier PCS:

- 32 kb/s ADPCM speech encoding in the near future to take advantage of the low complexity and low power consumption, and to provide low-delay high-quality speech.
- Flexible radio link architecture that will support multiple data rates from several kb/s to several hundred kb/s. This is needed to permit evolution in the future to lower bit rate speech as technology improvements permit high quality without excessive power consumption or transmission delay, and to provide multiple data rates for data transmission and messaging.
- Low transmitter power (≤ 25 mW average) with adaptive power control to maximize talk time and data transmission time. This incurs short radio range which requires many base stations to cover a large region. Thus, base stations must be small and inexpensive, like cordless telephone phone points or the Metricom wireless data base stations.
- Low complexity signal processing to minimize power consumption. Complexity one-tenth that of digital cellular or high-tier PCS technologies is required [1]. With only several tens of milliwatts (or less under power control) required for transmitter power, signal processing power becomes significant.
- Low co-channel interference and high coverage area design criteria. In order to provide high-quality service over a large region, at least 99 percent of any covered area must receive good or better coverage, and be below acceptable co channel interference limits. This implies less than 1 percent of a region will receive marginal service. This is an order-of-magnitude higher service requirement than the ten percent of a region permitted to receive marginal service in vehicular cellular system (high-tier PCS) design criteria.
- Four-level phase modulation with coherent detection to maximize radio link performance and capacity with low complexity.
- Frequency division duplexing to relax the requirement for synchronizing base station transmissions over a large region.

Such technologies and systems have been designed, prototyped, and laboratory-and field-tested and evaluated for several years [1, 2, 7, 16-23]. The viewpoint expressed here is consistent with the progress in the Joint Technical Committee (JTC) of the U.S. standards bodies, Telecommunications Industry Association (TIA) and Committee T1 of the Alliance for Telecommunications Industry Solutions (ATIS). Many technologies and systems were submitted to the JTC for consideration for wireless PCS in the new 1.9 GHz frequency bands for use in the United States [12] Essentially all of the technologies and systems listed in Table 1, and some others, were submitted in late 1993. It was evident that there were at least two, and perhaps three distinct different classes of submissions. No systems optimized for packet data were submitted, but some of the technologies are optimized for voice.

One class of submissions was the group labeled High Power Systems, Digital Cellular (High-Tier PCS) in Table 1. These are the technologies discussed previously in this article. They are highly optimized for low bit-rate voice, and therefore have somewhat limited capability for serving packet-data applications. Since it is clear that wireless services to wide ranging high speed mobiles will continue to be needed, and that the technology described above for low-tier PCS may not be optimum for such services, Fig. 1 shows a continuing evolution and need in the future for high-tier PCS systems that are the equivalent of today's cellular radio. There are more than 100 million vehicles in the United States alone. In the future, most, if not all, of these will be equipped with high-tier cellular mobile phones. Therefore, there will be a continuing and rapidly expanding market for high-tier systems.

It is not clear what the future roles are for paging/messaging, cordless-telephone appliances, or wide area packet-data networks in an environment having widespread contiguous coverage by low-tier and high-tier PCS.

Another class of submissions to the JTC [12] included the Japanese Personal Handiphone System (PHS), and a technology and system originally developed at Bellcore, but carried forward to prototypes, and submitted to the JTC, by Motorola and Hughes Network Systems. This system was known as Wireless Access Communications System (WACS).[2] These two submissions were so similar in their design objectives and system characteristics that, with the agreement of the delegations from Japan and the United States, the PHS and WACS submissions were combined under a new name, Personal Access Communication Systems (PACS), that was to incorporate the best features of both. This advanced, low-power wireless access system, PACS, was to be know as low-tier PCS. Both WACS/PACS and Handiphone (PHS) are shown in Table 1 as Low-Tier PCS and represent the evolution to low-tier PCS, on Fig. 1. The WACS/PACS/

[2] *WACS was known previously as Universal Digital Portable Communications (UDPC).*

87

Parameter	Cellular (high tier)	Low tier PCS	Capacity factor
Speech coding	8 kb/s (MOS 3.4) No tandem coding	32 kb/s (MOS 4.1) 3 or 4 tandem	x 4
Speech activity	Yes (MOS 3.2)	No (MOS 4.1)	x 2.5
Percentage of good areas	90%	99%	x 2
Propagation σ	8 dB	10 dB	x 1.5
Total: trading quality for capacity			x 30

■ **Table 4.** *A comparison of cellular (IS-54/IS-95) and low tier PCS (WACS/PACS). Capacity comparisons made without regard to quality factors, complexity, and cost per base station are not meaningful.*

UDPC system and technology are discussed in [1, 2, 16-23].

In the JTC, submissions for PCS of DECT and CT-2 and their variations were also lumped under the class of low-tier PCS, even though these advanced digital cordless telephone technologies were somewhat more limited in their ability to serve all of the low-tier PCS needs. They are included under Digital Cordless technologies in Table 1. Other technologies and systems were also submitted to the JTC for high-tier and low-tier applications, but they have not received widespread industry support.

One wireless access application discussed earlier that is not addressed by either high-tier or low-tier PCS is the high-speed WLAN application. Specialized high-speed WLANs also are likely to find a place in the future. Therefore, their evolution is also continued in Fig. 1. The figure also recognizes that widespread low-tier PCS can support data at several hundred kb/s, and thus can satisfy many of the needs of WLAN users.

It is not clear what the future roles are for paging/messaging, cordless telephone appliances, or wide area packet-data networks in an environment with widespread contiguous coverage by low-tier and high-tier PCS. Thus, their extensions into the future are indicated with a (?) on Fig. 1.

Those who may object to the separation of Wireless PCS into high tier and low tier, should review this section again, and note that we have two tiers of PCS now. On the voice side there is Cellular Radio, i.e., high-tier PCS, and cordless telephone, i.e., an early form of low-tier PCS. On the data side there is wide area data, i.e., high-tier data PCS, and WLANs, i.e., perhaps a form of low-tier data PCS. In their evolutions, these all have the trajectories discussed and shown on Fig. 1 that point surely toward low-tier PCS. It is this low-tier PCS that marketing studies continue to project is wanted by more than half the U.S. households or by half of the people, a potential market of over 100 million subscribers in the United States alone. Similar projections have been made worldwide.

Quality, Capacity, and Economic Issues

Although the several trajectories toward low-tier PCS discussed in the previous section are clear, it does not fit the existing wireless communications paradigms. Thus, low-tier PCS has attracted less attention than the systems and technologies that are compatible with the existing paradigms. Some examples are cited in the following paragraphs.

The need for intense interaction with an intelligent network infrastructure in order to manage mobility is not compatible with the cordless telephone appliance paradigm. In that paradigm, independence of network intelligence, and base units that mimic wireline telephones, are paramount.

Wireless data systems often do not admit to the dominance of wireless voice communications, and, thus, do not take advantage of the economics of sharing network infrastructure and base station equipment. Also, wireless voice systems often do not recognize the importance of data and messaging, and, thus, only add them in as "bandaids" to systems.

The need for a dense collection of many low-complexity low-cost low-tier PCS base stations interconnected with inexpensive fixed-network facilities (copper or fiber based) does not fit the cellular high-tier paradigm that expects sparsely distributed $1 million cell sites. Also, the need for high transmission quality to compete with wireline telephones is not compatible with the drive toward maximizing users-per-cell-site and per MHz to minimize the number of expensive cell sites. These concerns, of course, ignore the hallmark of frequency-reusing cellular systems. That hallmark is the production of almost unlimited overall system capacity by reducing the separation between base stations.

This list could be extended, but the above examples are sufficient, along with the earlier sections of the paper, to indicate the many complex interactions among circuit quality, spectrum utilization, complexity (circuit and network), system capacity, and economics that are involved in the design compromises for a large, high-capacity wireless-access system. Unfortunately, the tendency has been to ignore many of the issues, and focus on only one, e.g., the focus on cell site capacity that drove the development of digital-cellular high-tier systems in the United States. Interactions among circuit quality, complexity, capacity and economics are considered in the following sections.

Capacity, Quality, and Complexity

Although "capacity" comparisons frequently are made without regard to circuit quality, complexity, or cost per base station, such comparisons are not meaningful. An example in Table 4 compares capacity factors for U.S. cellular or high-tier PCS technologies with the low-tier PCS technology, PACS/WACS. The Mean Opinion Scores (MOS) (noted in Table 4) for speech coding are discussed in reference [1]. Detection of speech activity and turning off the transmitter during times of no activity is implemented in IS-95. Its impact on MOS also is noted in reference [1]. A similar technique has been proposed as E-TDMA for use with IS-54, and is discussed with respect to TDMA systems in reference [1]. Note that the use of low bit-rate speech coding combined with speech activity degrades the high-tier system's quality by nearly one full MOS point on the 5-point MOS scale when compared to 32 kb/s ADPCM. Tandem encoding is discussed in the previous section. These speech

quality-degrading factors alone provide a base station capacity increasing factor of x 4 x 2.5 = x 10 over the high speech-quality low-tier system! Speech coding, of course, directly affects base station capacity and thus overall system capacity by its effect on the number of speech channels that can fit into a given bandwidth.

The allowance of extra system margin to provide coverage of 99 percent of an area for low-tier PCS *versus* 90 percent coverage for high-tier is discussed in the previous section and [1]. This additional quality factor costs a capacity factor of x 2. The last item in Table 4 does not change the actual system, but only changes the way that frequency reuse is calculated. The additional 2-dB margin in standard deviation, σ, allowed for coverage into houses and small buildings for low-tier PCS, costs yet another factor of x 1.5 in calculation only. Frequency reuse factors affect the number of sets of frequencies required, and thus the bandwidth available for use at each base station. Thus, these factors also affect the base station capacity and the overall system capacity.

For the example in Table 4, significant speech and coverage quality has been traded for a factor of x 30 in base station capacity!! While base station capacity affects overall system capacity directly, it should be remembered that overall system capacity can be increased arbitrarily by decreasing the spacing between base stations. Thus, if the PACS low-tier PCS technology were to start with a base station capacity of x 0.5 of AMPS cellular[3] (a much lower figure than the x 0.8 sometimes quoted [12]), and then were degraded in quality as described above to yield the x 30 capacity factor, it would have a resulting capacity of x 15 of AMPS! Thus, it is obvious that making such a base station capacity comparison without including quality is not meaningful.

Economics, System Capacity, and Coverage Area Size

Claims are sometimes made that low-tier PCS cannot be provided economically, even though IT is what the user wants. These claims are often made based on economic estimates from the "cellular paradigm." These include:

- Very low estimates of market penetration, much less than cordless telephones, and often even less than cellular.
- High estimates of base station costs more appropriate to high-complexity high-cost cellular technology than to low-complexity low-cost low-tier technology.

Such economic estimates are often done by making "absolute" economic calculations based on very uncertain input data. The resulting estimates for low-tier and high-tier are often closer together than the large uncertainties in the input data. A perhaps more realistic approach for comparing such systems is to vary only one or two parameters while holding all others fixed, and then look at relative economics between high-tier and low-tier systems. This is the approach used in the following examples.

Example 1 – In the first example (see textbox), the number of channels per MHz is held constant for cellular and for low-tier PCS. Only the spacing is varied between base stations, e.g., cell sites for cellular and radio ports for low-tier PCS, to account for the differences in transmitter power, antenna height, etc. In this example, overall system capacity varies directly as the square of base station spacing, but base station capacity is the same for both cellular and low-tier PCS. For the typical values in the example, the resulting low-tier system capacity is x 400 greater, only because of the closer base station spacing. If the two systems were to cost the same, the equivalent low-tier PCS base stations would have to cost less than $2,500.

This cost is well within the range of estimates for such base stations, including equivalent infrastructure. These low-tier PCS base stations are of comparable or lower complexity than cellular vehicular subscriber sets, and large-scale manufacture will be needed to produce the millions that will be required. Also, land, building, antenna tower and legal fees for zoning approval, or rental of expensive space on top of commercial buildings, represent large expenses for cellular cell sites. Low-tier PCS base stations that are mounted on utility poles and sides of buildings will not incur such large additional expenses. Therefore, costs of the order of magnitude indicated above seem reasonable in large quantities. Note that, with these estimates, the per-wireless-circuit cost of the low-tier PCS circuits would be only $14/circuit compared to $5,555/circuit for the high-tier circuits. Even if there were a factor of 10 error in cost estimates, or a reduction of channels per radio port of a factor of 10, the per-circuit cost of low-tier PCS would

System Capacity/Coverage Area Size/Economics

Example 1

Assume channels/MHz are the same for cellular and PCS
Cell site: spacing = 20,000 ft cost = $1 M
PCS port: spacing = 1,000 ft
PCS system capacity is $(20000/1000)^2$ = 400 x cellular capacity

Then, for the system costs to be the same
Port cost = ($1M/400) = $2,500, a reasonable figure

If, cell site and port each have 180 channels
Cellular cost/circuit = $1M/180 = $5,555/circuit
PCS cost/circuit = $2500/180 = $14/circuit

Example 2

Assume equal cellular and PCS system capacity
Cell site: spacing = 20,000 ft
PCS port: spacing = 1,000 ft
If, a cell site has 180 channels

Then, for equal system capacity, a PCS port needs 180/400 < 1 channel/port!!

Example 3

A Quality/cost trade
Cell site: Spacing = 20,000 ft cost = $1 M channels = 180
PCS port: Spacing = 1,000 ft cost = $2,500

Cellular to PCS: base station spacing capacity factor = x 400
PCS to Cellular "quality" reduction factors:

32 kb/s to 8 kb/s speech	x 4
Voice activity (buying)	x 2
99% to 90% good areas	x 2
Both in same environment (same σ)	x 1
capacity factor traded	x 16

180 ch/16 = 11.25 channels/port then, $2500/11.25 = $222/circuit
and remaining is x 400/16 = x 25 system capacity of PCS over cellular

[3] Note that the x 0.5 factor is an arbitrary factor taken for illustrating this example. The so called x AMPS factors are only with regard to base station capacity, although they are often misused as system capacity.

still be only $140/circuit, which is still much less than the per-circuit cost of high-tier.

Example 2 – In the second example (see textbox), the overall system capacity is held constant, and the number of channels/port, i.e., channels/(base station) is varied. In this example, less than 1/2 channel/port is needed, again indicating the tremendous capacity that can be produced with close-spaced low-complexity base stations.

Example 3 – Since the first two examples are somewhat extreme, the third example (see textbox), uses a more moderate, intermediate approach. In this example, some of the cellular high-tier channels/(base station) are traded to yield higher quality low-tier PCS as in the previous subsection. This reduces the channels/port to 11 +, with an accompanying increase in cost/circuit up to $222/circuit, which is still much less than the $5,555/circuit for the high-tier system. Note, also, that the low-tier system still has x 25 the capacity of the high-tier system!

Low-tier base station (PORT) cost would have to exceed $62,500 for the low-tier per-circuit cost to exceed that of the high-tier cellular system. Such a high port cost far exceeds any existing realistic estimate of low-tier system costs.

It can be seen from these examples, and particularly Example 3, that the circuit economics of low-tier PCS are significantly better than for high-tier PCS, IF the user demand and density is sufficient to make use of the large system capacity. Considering the high penetration of cordless telephones, the rapid growth of cellular handsets, and the enormous market projections for "wireless PCS" noted earlier in this paper, filling such high capacity in the future would appear to be certain. The major problem is providing rapidly the widespread coverage (buildout) required by the FCC in the United States. If this unrealistic regulatory demand can be overcome, low-tier wireless PCS promises to provide *the* wireless personal communications that everyone wants.

With the continuing problems and delays in initial deployments, there is increasing concern throughout the industry as to whether CDMA is a viable technology for high capacity cellular applications.

Other Issues

Several issues in addition to those addressed in the pervious two sections continue to be raised with respect to low-tier PCS. These are treated in this section.

Improvement of Batteries

Frequently, the suggestion is made that battery technology will improve so that high-power handsets will be able to provide the desired five or six hours of talk time in addition to 10 or 12 hours of standby time, and still weigh less than half of the weight of today's smallest cellular handset batteries. This "hope" does not take into account the maturity of battery technology, and the long history (many decades) of concerted attempts to improve it. Increases in battery capacity have come in small increments, a few percent, and very slowly over many years, and the shortfall is well over a factor of 10. In contrast, integrated electronics and radio frequency devices needed for low-power low-tier PCS continue to improve and to decrease in cost by factors of greater than 2 in time spans on the order of a year or so. It also should be noted that, as the energy density of a battery is increased, the energy release rate per volume must also increase in order to supply the same amount of power. If energy storage density and release rate are increased significantly, the difference between a battery and a bomb become indistinguishable! The likelihood of a x 10 improvement in battery capacity appears to be essentially zero. If even a modest improvement in battery capacity were possible, many people would be driving electric vehicles.

New Technology

New technology, e.g., spread spectrum or CDMA, is sometimes offered as a solution to both the high-tier cell site capacity and transmitter power issues. However, as these new technologies are pursued vigorously, it becomes increasingly evident that the early projections were considerably over-optimistic, that the base station capacity will be about the same as other technologies [1], and that the high complexity will result in more, not less, power consumption.

With the continuing problems and delays in initial deployments, there is increasing concern throughout the industry as to whether CDMA is a viable technology for high capacity cellular applications. With the passage of time, it is becoming more obvious that Viterbi was correct in his 1985 paper in which he questioned the use of spread spectrum for commercial communications [33].

Thus, it is clear that new high-complexity high-tier technology will not be a substitute for low-complexity, low-power low-tier PCS.

People Only Want One Handset

This issue is often raised in support of high-tier cellular handsets over low-tier handsets. While the statement is likely true, the assumption that *the* handset must work with high-tier cellular is not. Such a statement follows from the current large usage of cellular handsets; but such usage results because that is the only form of widespread wireless service currently available, not because it is what people want. The statement assumes inadequate coverage of a region by low-tier PCS, and that low-tier handsets will not work in vehicles. The only way that high-tier handsets could serve the desires of people discussed earlier would be for an unlikely "breakthrough" in battery technology to occur [7]. However, a low-tier system can cover economically any large region having some people in it. (It will not cover rural or isolated areas — but, by definition, there is essentially no one there to want communications anyway).

Low-tier handsets will work in vehicles on village and city streets at speeds up to 30 or 40 miles per hour, and the required handoffs make use of computer technology that is rapidly becoming inexpensive. Highways between populated areas, and also streets within them, will need to be covered by high-tier cellular PCS, but, users are likely

to use vehicular sets in these cellular systems. Frequently the vehicular mobile user will want a different communications device anyway, e.g., a hands-free phone. The use of hands-free phones in vehicles is becoming a legal requirement in some places now, and is likely to become a requirement in many more places in the future. Thus, handsets may not be legally usable in vehicles anyway. With widespread deployment of low-tier PCS systems, *the one* handset of choice will be the low-power low-tier PCS pocket handset or voice/data communicator.

There are approaches for integrating low-tier pocket phones or pocket communicators with high-tier vehicular cellular mobile telephones. The user's identity could be contained either in memory in the low-tier set, or in a small smart card inserted into the set, as is a feature of the European GSM system. When entering an automobile, the small low-tier communicator or card could be inserted into a receptacle in a high-tier vehicular cellular set installed in the automobile.[4] The user's identity would then be transferred to the mobile set. The mobile set could then initiate a data exchange with the high-tier system, indicating that the user could now receive calls at that mobile set. This information about the user's location would then be exchanged between the network intelligence so that calls to the user could be correctly routed.[5] In this approach the radio sets are optimized for their specific environments, high-power high-tier vehicular or low-power low-tier pedestrian, as discussed earlier, and the network access and call routing is coordinated by the interworking of network intelligence. This approach does not compromise the design of either radio set or radio system. It places the burden on network intelligence technology that benefits from the large and rapid advances in computer technology.

The approach of using different communications devices for pedestrians than for vehicles is consistent with what has actually happened in other applications of technology in similarly different environments. For example, consider the case of audio cassette tape players. Pedestrians often carry and listen to small portable tape players with lightweight headsets (e.g., a Walkman[6]). When one of these people enters an automobile, he or she often removes the tape from the Walkman and inserts it into a tape player installed in the automobile. The automobile player has speakers that fill the car with sound. The Walkman is optimized for a pedestrian, whereas the vehicular-mounted player is optimized for an automobile. Both use the same tape, but they have separate tape heads, tape transports, audio preamps, etc. They do not attempt to share electronics. In this example, the tape cassette is the information-carrying entity similar to the user identification in the personal communications example discussed earlier. The main points are that the information is shared among different devices, but the devices are optimized for their environments and do not share electronics.

Similarly, a high-tier vehicular-cellular set does not need to share oscillators, synthesizers, signal processing, or even frequency bands or protocols with a low-power low-tier pocket-size communicator. Only the information identifying the user and where he or she can be reached needs to be shared among the intelligence elements, e.g., routing logic, databases, and common channel signaling [1, 22] of the infrastructure networks. This information exchange between network intelligence functions can be standardized and coordinated among infrastructure subnetworks owned and operated by different business entities (e.g., vehicular cellular mobile radio networks, and intelligent low-tier PCS networks). Such standardization and coordination are the same as are required today to pass intelligence among local exchange networks and interexchange carrier networks.

The approach of using different communications devices for pedestrians than for vehicles is consistent with what has actually happened in other applications of technology in similarly different environments.

Other Environments – Low-tier personal communications can be provided to occupants of airplanes, trains, and buses by installing compatible low-tier radio access ports inside these vehicles. The ports can be connected to high-power high-tier vehicular cellular mobile sets or to special air-ground or satellite-based mobile communications sets. Intelligence between the internal ports and mobile sets could interact with cellular mobile, air-ground, or satellite networks in one direction, using protocols and spectrum allocated for that purpose, and with low-tier personal communicators in the other direction to exchange user identification and route calls to and from users inside these large vehicles. Radio isolation between the low-power units inside the large metal vehicles and low-power systems outside the vehicles can be ensured by using windows that are opaque to the radio frequencies. Such an approach also has been considered for automobiles (i.e., a radio port for low-tier personal communications connected to a cellular mobile set in a vehicle so that the low-tier personal communicator can access a high-tier cellular network. This could be done in the United States using unlicensed PCS frequencies within the vehicle.)

High-Tier to Low-Tier or Low-Tier to High-Tier Dual Mode

Industry and the FCC in the United States appear willing to embrace multi-mode handsets for operating in very different high-tier cellular systems, e.g., analog FM AMPS, TDMA IS-54, and CDMA IS-95. Such sets incur significant penalties for dual mode operation with dissimilar air interface standards, and, of course, incur the high-tier complexity penalties.

It has been suggested that multi-mode high-tier and low-tier handsets could be built around one air-interface standard, for example, TDMA IS-54 or GSM. When closely spaced low-power base stations were available, the handset could "turn off" unneeded power-consuming circuitry, e.g., the multipath equalizer. The problem with this approach is that the handset is still encumbered with power-consuming and quality-reducing signal processing inherent in the high-tier technology,

[4] Inserting the small personal communicator in the vehicular set would also facilitate charging the personal communicator's battery.

[5] This is a feature proposed for FPLMTS in CCIR Rec. 687.

[6] Walkman is a registered trademark of Sony Corporation.

e.g., error correction decoding, and low-bit-rate speech encoding and decoding.

An alternative "dual-mode" low-tier, high-tier system based on a common air-interface standard can be configured around the low-tier PACS/WACS system, if such a dual-mode system is deemed desirable in spite of the discussion in this article. The range of PACS can readily be extended by increasing transmitter power and/or the height and gain of base station antennas. With increased range, the multipath delay-spread will be more severe in some locations [24-26]. Two different solutions to the increased delay-spread can be employed, one for the downlink and another for the uplink. The PACS radio-link architecture has a specified bit sequence, i.e., a unique word, between each data word on the TDM downlink [16, 17]. This unique word can be used as a training sequence for setting the tap weights of a conventional equalizer added to subscriber sets for use in a "high-tier" PACS mode. Since received data can be stored digitally [27, 28], tap weights can be trimmed, if necessary, by additional "passes" through an adaptive equalizer algorithm, e.g., a decision feedback equalizer algorithm.

The signaling, control processing, and data base interactions required for wireless access PCS are considerably greater than those required for fixed place-to-place networks, but that fact must be accepted when considering such networks.

The PACS TDMA uplink has no "unique word." However, the "high-tier" uplink will terminate on a base station that can support greater complexity, but still be no more complex than the high-tier cellular technologies. Research at Stanford University has indicated that blind equalization, using constant-modulus algorithms (CMA) [29, 30], can be effective for equalizing the PACS uplink. Techniques have been developed for converging the CMA equalizer on the short TDMA data burst.

Advantages of building a dual-mode high-tier, low-tier PCS system around the low-tier PACS air-interface standard are that:

- The interface can still support small low-complexity, low-power, high-speech-quality low-tier handsets.
- Both data and voice can be supported in a PACS personal communicator.
- In high-tier low-tier dual mode PACS sets, circuits used for low-tier operation will also be used for high-tier operation, with additional circuits being activated only for high-tier operation.
- The flexibility built into the PACS radio link to handle different data rates from 8 kb/s to several hundred kb/s will be available to both modes of operation.

Infrastructure Networks

It is beyond the scope of this article to consider the details of PCS network infrastructures. However, there are perhaps as many network issues as there are wireless access issues discussed herein [22, 23, 31, 32]. With the possible exception of the self-organizing WLANS, wireless PCS technologies serve as access technologies to large integrated intelligent fixed communications infrastructure networks.

These infrastructure networks must incorporate intelligence i.e., data-base storage, signaling, processing and protocols, to handle both small-scale mobility, i.e., handoff from base station to base station as users move, and large-scale mobility, i.e., providing service to users who roam over large distances, and perhaps from one network to another. The fixed infrastructure networks also must provide the interconnection among base stations and other network entities, e.g., switches, data bases, and control processors. Of course, existing cellular mobile networks now contain or are incorporating these infrastructure network capabilities. However, existing cellular networks are small compared to the expected size of future high-tier and low-tier PCS networks, e.g., 20 million cellular users in the United States compared with perhaps 100 million users or more each in the future for high-tier and low-tier PCS.

Several other existing networks have some of the capabilities needed to serve as access networks for PCS. Existing networks that could provide fixed base station interconnection include:

- Local exchange networks that could provide interconnection using copper or glass-fiber distribution facilities.
- Cable TV networks that could provide interconnection using new glass-fiber and coaxial-cable distribution facilities.
- Metropolitan fiber digital networks that could provide interconnection in some cities in which they are being deployed.

Networks that contain intelligence, e.g., databases, control processors, and signaling that is suitable, or could be readily adapted, to support PCS access include:

- Local exchange networks that are equipped with signaling system 7 common channel signaling (SS7 CCS), data bases and digital control processors.
- Interexchange networks that are similarly equipped.

Data networks, e.g., the Internet, could perhaps be adapted to provide the needed intelligence for wireless data access, but it does not have the capacity needed to support large voice/data wireless low-tier PCS access.

Many entities and standards bodies worldwide are working on the access network aspects of wireless PCS. The signaling, control processing, and data base interactions required for wireless access PCS are considerably greater than those required for fixed place-to-place networks, but that fact must be accepted when considering such networks.

Low-tier PCS, when viewed from a cellular high-tier paradigm, requires much greater fixed interconnection for the much closer spaced base stations. However, when viewed from a cordless telephone paradigm of a base unit for every handset, and perhaps several base units per wireline, the requirement is much less fixed interconnection because of the concentration of users and trunking that occurs at the multi-user base stations. One should remember that there are economical fixed wireline connections to almost all houses and business offices in the United States now. If wireless access displaces some of the wireline

connections, as expected, the overall need for fixed interconnection could decrease!

Conclusion

Wireless personal communications embraces about seven relatively distinct groups of tetherless voice and data applications or services having different degrees of mobility for operation in different environments. Many different technologies and systems are evolving to provide the different perceived needs of different groups. Different design compromises are evident in the different technologies and systems. The evidence suggests that the evolutionary trajectories are aimed toward at least three large groups of applications or services, namely, high-tier PCS (current cellular radio), high-speed wireless local-area networks (WLANS), and low-tier PCS (an evolution from several of the current groups). It is not clear to what extent several groups, e.g., cordless telephones, paging, and wide area data, will remain after some merging with the three large groups. Major considerations that separate current cellular technologies from evolving low-tier low-power PCS technologies are speech quality, complexity, flexibility of radio-link architecture, economics for serving high-user-density or low-user-density areas, and power consumption in pocket carried handsets or communicators. High-tier technologies make use of large complex expensive cell sites and have attempted to increase capacity and reduce circuit costs by increasing the capacity of the expensive cell sites. Low-tier technologies increase capacity by reducing the spacing between base stations, and achieve low circuit cost by using low-complexity low-cost base stations. The differences between these approaches result in significantly different compromises in circuit quality and power consumption in pocket sized handsets or communicators. These kinds of differences also can be seen in evolving wireless systems optimized for data. Advantages of the low-tier PACS/WACS technology are reviewed in the article, along with techniques for using that technology in high-tier PCS systems.

References

[1] D. C. Cox, "Wireless Network Access for Personal Communications," *IEEE Commun. Mag.*, Dec. 1992, pp. 96-115.
[2] J. E. Padgett, T. Hattori and C. Gunther, "Overview of Wireless Personal Communications," *IEEE Commun. Mag.*, Jan. 1995, pp. 28-41.
[3] D. J. Goodman, "Trends in Cellular and Cordless Communications," *IEEE Commun. Mag.*, June 1991, pp. 31-40.
[4] R. Steele, "Deploying Personal Communications Networks," *IEEE Commun. Mag.*, Sept. 1990, pp. 12-15.
[5] *IEEE Commun. Mag.*, Special Issue on "Wireless Personal Communications," Jan. 1995.
[6] R. Schneideman, "Spread Spectrum Gains Wireless Applications," *Microwaves and RF*, May 1992, pp. 31-42.
[7] D. C. Cox, "Personal Communications-A Viewpoint," *IEEE Commun. Mag.*, Nov. 1990, pp. 8-20.
[8] A. Wong, "Regulating Public Wireless Networks," Workshop on Lightwave, Wireless and Networking Technologies, the Chinese University of Hong Kong, Hong Kong, Aug. 24, 1994.
[9] J. C-I, Chuang, "Performance Issues and Algorithms for Dynamic Channel Assignment," *IEEE JSAC*, Aug. 1993.
[10] J. C-I Chuang, N.R. Sollenberger, and D.C. Cox, "A Pilot Based Dynamic Channel Assignment Scheme for Wireless Access TDMA/FDMA Systems," Proc. IEEE ICUPC'93, Ottawa, Canada, Oct. 12-15, 1993, pp. 706-712, and *Int'l J. of Wireless Info. Networks*, vol. 1, no. 1, Jan. 1994, pp. 37-48.
[11] J. C-I, Chuang, "Performance Limitations of TDD Wireless Personal Communications with Asynchronous Radio Ports," *Electronic Letters*, vol. 28, March 12, 1992, pp. 532-533.
[12] C. I. Cook, "Development of Air Interface Standards for PCS," *IEEE Personal Commun.*, Fourth Quarter, 1994, pp. 30-34.
[13] Bell System Technical Journal (BSTJ), Special Issue on Advanced Mobile Phone Service (AMPS), vol. 58, Jan. 1979.
[14] R. H. Katz, "Adaptation and Mobility in Wireless Information Systems," *IEEE Personal Commun.*, First Quarter, 1994, pp. 6-17.
[15] D. G. Steer, "Coexistence and Access Etiquette in the United States Unlicensed PCS Band," *IEEE Personal Communications Magazine*, Forth Quarter, 1994, pp. 36-43.
[16] D. C. Cox, "A Radio System Proposal for Widespread Low-power Tetherless Communications," *IEEE Trans. on Commun.*, Feb. 1991, pp. 324-335.
[17] Bellcore Technical Advisories, "Generic Framework Criteria for Universal Digital Personal Communications Systems (PCS)," FA-TSY-001013, Issue 1, March 1990 and FA-NWT-001013, Issue 2, Dec. 1990, and Technical Reference, "Generic Criteria for Version 0.1 Wireless Access Communications Systems (WACS), Issue 1, Oct. 1993 and Revision 1, June 1994.
[18] D.C. Cox, "Universal Portable Radio Communications," *IEEE Trans. on Veh. Tech.*, Aug. 1985, pp. 117-121.
[19] D. C. Cox, H. W. Arnold, and P.T. Porter, "Universal Digital Portable Communications-A System Perspective," *IEEE JSAC*, Vol. JSAC-5 pp. 764-773, June 1987.
[20] D. C. Cox, "Research Toward a Wireless Digital loop," *Bellcore Exchange*, vol. 2, Nov./Dec. 1986, pp. 2-7.
[21] D. C. Cox, "Universal Digital Portable Radio Communications," Proc. IEEE, vol. 75, April 1987, pp. 436-477.
[22] D. C. Cox, "Portable Digital Radio Communications-An Approach to Tetherless Access," *IEEE Commun. Mag.*, July 1989, pp. 30-40.
[23] D. C. Cox, W. G. Gifford, and H. Sherry, "Low-Power Digital Radio as a Ubiquitous Subscriber Loop," *IEEE Commun. Mag.*, March 1991, pp. 92-95.
[24] D. M. J. Devasirvatham, "Radio Propagation Studies in a Small City for Universal Portable Communications," Conference Record, IEEE VTC'88, Philadelphia, PA, June 15-17, 1988, pp. 100-104.
[25] D. C. Cox, "Delay-Doppler Characteristics of Multipath Propagation at 910 MHz in a Suburban Mobile Radio Environment," *IEEE Trans. on Antennas and Propagation*, Sep. 1972, pp. 625-635.
[26] D. C. Cox, "Multipath Delay Spread and Path Loss Correlation for 910 MHz Urban Mobile Radio Propagation," *IEEE Trans. on Vehicular Technology*, Nov. 1977, pp. 340-344.
[27] N. R. Sollenberger and J. C-I. Chuang, "Low-Overhead Symbol Timing and Carrier Recovery for TDMA Portable Radio Systems," *IEEE Trans. on Commun.* Oct. 1990, pp. 1886-1892,.
[28] N.R. Sollenberger, "An Experimental VLSI Implementation of Low-Overhead Symbol Timing and Frequency Offset Estimation for TDMA Portable Radio Applications," IEEE GLOBECOM '90, San Diego, CA, Dec. 2-5, 1990, pp. 1701-1711.
[29] J. R. Treichler and B. G. Agee, "A New Approach to Multipath Correction of Constant Modulus Signals," *IEEE Trans. on Acoustics, Speech and Signal Processing*, April 1983, pp. 459-472.
[30] Y. Sato, "A Method of Self-Recovering Equalization for Multilevel Amplitude Modulation Systems," *IEEE Trans. on Commun.*, June 1975, pp. 679-682.
[31] B. Jabbari *et al.*, "Network Issues for Wireless Personal Communications," *IEEE Commun. Mag.*, Jan. 1995, pp. 88-98.
[32] M. Zaid, "Personal Mobility in PCS," *IEEE Personal Commun.*, Fourth Quarter, 1994, pp. 12-16.
[33] A. J. Viterbi, "When Not to Spread Spectrum-A Sequel," *IEEE Commun. Mag.*, April 1985, pp. 12-17.

Wireless personal communications embraces about seven relatively distinct groups of tetherless voice and data applications or services having different degrees of mobility for operation in different environments. Many different technologies and systems are evolving to provide the different perceived needs of different groups.

Biography

DONALD C. COX [F '79] did research at Bell Laboratories from 1968 to 1973 on advanced mobile radio systems that still provides basic input to the design of digital cellular, cordless, and PCS systems. From 1978 to 1993 he led and was actively involved in pioneering wireless research, first at Bell Labs and then at Bellcore, that started and fueled the current explosion in wireless personal communications. He was instrumental in evolving this research into the WACS/PACS specification being standardized by the U.S. TIA/T1 JTC. For this pioneering work, he received the IEEE 1993 Alexander Graham Bell Medal and Bellcore Fellow Award, and was elected into the National Academy of Engineering. He received the IEEE Morris E. Leeds award in 1985 and the Prize Guglielmo Marconi from Italy in 1983 and is a Fellow of the AAAS and the RCA. He holds 12 patents, has authored or coauthored more than 75 journal papers, including three that won prizes, was coauthor of a book, *Microwave Mobile Communications*, and has been guest editor of special issues on Wireless Communications in IEEE journals. He managed Radio Research at Bellcore from 1984 to 1993, and is currently the Harald Trap Friis Professor of Engineering and director of the Telecommunications Center at Stanford University, Stanford, California. He received B.S. and M.S. degrees in electrical engineering and an Honorary Dr. of Science from the University of Nebraska, and a Ph.D. in electrical engineering from Stanford.