

Wireless Digital Communication: A View Based on Three Lessons Learned

Shannon's information theory is applied to PCN, with examples of improvements in satellite communication, magnetic recording and modem technologies that have approached information-theoretic limits.

Andrew J. Viterbi

The first two installments of "Technologies on the Horizon" in September and November, 1990, featured personal viewpoints by Ray Steele and Don Cox on the emerging area of personal communication networks (PCNs). Both authors addressed key issues that need to be resolved to expedite large scale deployment of low power, low cost, low complexity wireless access technologies.

My request for comments or reactions from members of the communications community drew a passionate and informed response from a third leading contributor to recent advances in personal communications: Andrew J. Viterbi, Chief Technical Officer of QUALCOMM, Inc., took exception to Don Cox's view that variations in power control and soft hand-off imperfections will cause loss of system capacity in PCNs based on spread-spectrum techniques.

As Viterbi noted in his response to me, "... Cox faults spread-spectrum digital techniques for excessive complexity, but at the same time claims they will be ineffective in 'real-world' situations." Viterbi then adds, "If he were right on the latter, there would be no need to argue the former. He is, I contend, wrong on both, because the complexity, which is exclusively in the digital processing, can be made as cheap and virtually as power thrifty, as needed by market-demand influence on economies of scale."

At my invitation, Andy Viterbi contributes his views this month on PCN design and complexity to inform readers of IEEE Communications Magazine about recent advances in spread-spectrum approaches to wireless digital communications. Viterbi outlines the implications of Shannon's information theory for PCN design, providing examples of improvements in satellite communication, magnetic recording and modem technologies that have increasingly approached information-theoretic limits. He emphasizes the application of Shannon's work to control of four types of interference: multi-user, multiple cell, multipath and multiple media. Viterbi then closes by applying "Shannon's third lesson" to a comparison of code-division multiple access (CDMA), time-division multiple access (TDMA) and frequency-division multiple access (FDMA) techniques in personal communications.

This month's column, together with earlier "Technologies on the Horizon" contributions by Ray

Steele and Don Cox, and references provided by all three authors, constitute a useful tutorial for interested, nonspecialist readers on the major technical alternatives currently being pursued as personal communication networks move from the laboratory to field experiments and initial service applications. Of course, the real feasibility test for any proposed approach to PCNs will ultimately be demonstrated in sizable field experiments involving large numbers of users, cells and very imperfect, "real-world" transmission environments. Such field experiments are currently under way, and we can expect to learn about real-world performance, complexity and cost in the next year or so, as field results for alternative wireless communications systems are reported.

With thanks to increasingly responsive and enthusiastic readership, I invite additional comments and suggestions for this column beyond the ones currently on my horizon. All correspondence should be sent to: Howard L. Lemberg, Bellcore, 445 South Street, Room 2M289, Morristown, NJ 07962-1910

It is a virtual certainty that wireline and fiber communication will be fully digital by the end of the century. It is almost as likely that the same will be true for most wireless communication. Beginning with broadcast of high quality digital audio, already underway in Japan and Europe, and with high definition digital video broadcasting on the horizon, it appears that by the end of the decade digital modulation will have superseded analog in virtually all forms of broadcasting. Digital communication provides excellent reproduction and greatest efficiency of transmission bandwidth and power through effective utilization of two fundamental techniques: source compression coding to greatly reduce the transmission rate for a given degree of fidelity; and error control coding to further reduce signal-to-noise and bandwidth requirements.

While digital modulation is also certain to become the universal choice for two-way personal and mobile communication, and while the two techniques just mentioned will be equally important here for the same reasons, there is another even

ANDREW VITERBI is Vice Chairman, and Chief Technical Officer, QUALCOMM, Inc.

Technologies on the Horizon is edited by Howard Lemberg of Bellcore.

Shannon's theory can be summarized in three lessons, which are transforming all communication products and services that surround us.

more critical need for such applications: to afford many users simultaneous and equal access to the network. Whichever multiple access technique is employed, the ultimate performance limitation is the system's susceptibility to interference. In narrowband multiple access techniques, whether frequency division or time division, this interference manifests itself as inter-user, intersymbol and multipath interference. Mitigation of the detrimental effects of these disturbances is limited to equalization and to isolation by frequency reuse with sufficient spatial separation. The first may be difficult to implement, particularly in mobile vehicles, while the second reduces spectrum efficiency.

Historical Precedents

Before delving deeper into these issues and the underlying lessons which lead to their resolution, it is useful to trace the origins of modern electronics from the discovery of the transistor in 1948. This singular event opened the solid state electronics era, which revolutionized business styles and lifestyles universally. The most visible fruits of the dramatic progress in solid state integration over the past two decades has been the ubiquitous application of computer power descending from large office applications to the most modest of home and personal users. A natural extension, with almost as great an impact, is to liberate the individual communicator from the tether, whether through cordless telephony, mobile cellular telephony or the emerging generic concept of the personal communication system. Equally natural and correct is the view that digital techniques, having evolved to such enormous speed and memory capabilities through computer applications of solid state technology (and to a lesser extent signal processing applications), will produce equally economical and powerful personal communication products.

The pitfall of the computer-communication analogy [1] is the exporting of not only hardware capabilities, but of system concepts as well, from the computer to the communication arena. Thus, for example, the concept of OSI layering, which is a logical extension of software hierarchies in computers, presents a deceptively over simplified view of communication systems and networks. The design of the lowest (or physical) layer has a much more profound impact on the feasibility of communication processes at higher layers than can be explained solely by issues of speed and memory. For an understanding of this key point, it is necessary to explore more fully the fundamentals of communications which by a remarkable coincidence were revealed in a research journal published in 1948, the same year in which the discovery of the transistor was made, and at the same research organization, the Bell Telephone Laboratories in Murray Hill, New Jersey.

Shannon's Three Lessons

This journal, "The Mathematical Theory of Communication" by Claude Shannon [2], more popularly referred to as Shannon's Information Theory, provided in a relatively abstract form the foundations for design of efficient wireless communication systems, including those seeking mul-

multiple access to a common medium. Of course it took longer than a decade before a sufficient number of communication engineers had been trained to understand and envision the ultimate practical embodiments of this theory, and about two decades beyond that before the above mentioned solid state evolution of the transistor reached the technological and economic level to implement these embodiments.

The essence of the Shannon theory, which pertains to digital communications, can be summarized in the form of three lessons, learned and applied over nearly a half century. Ranging in order from intuitive to somewhat surprising, these lessons are as follows:

- Never discard information prematurely that may be useful in making a decision until after all decisions related to that information have been completed.
- Completely separate techniques for digital source compression from those for channel transmission, even though the first removes redundancy and the second inserts it.
- In the presence of interference or jamming, intentional or otherwise, the communicator, through signal processing at both transmitter and receiver, can ensure that performance degradation due to the interference will be no worse than that caused by Gaussian noise at equivalent power levels. This implies that the jammer's optimal strategy is to produce Gaussian noise interference. Against such interference, the communicator's best waveform should statistically appear as Gaussian noise. Thus, the "minimax" solution to the contest is that signals and interference should all appear as noise which is as wideband as possible. This is a particularly satisfying solution when, as we shall see, one user's signal is another user's interference.

These three lessons are transforming all the communication products and services that surround us. The first lesson has radically changed the design approach to satellite communication over the past two decades. It has also had a major impact on the wireline modem industry, and is just beginning to be felt as an economic factor in the recording industry as well. Each field has applied Shannon's first lesson and assigned its own jargon to its implementation. In satellite communication they are called "forward error correction (FEC) with soft decision decoding," [3] In high quality wireline modems the techniques are known as "maximum-likelihood sequence estimation" [4] and "trellis coding." [5] In magnetic recording it is called "partial response maximum-likelihood (PRML) detection." [6]

More interesting is the impact of these techniques:

- In recording, they are replacing the very inefficient "peak detector" and increasing recording densities several fold.
- In wireline modems, adaptive linear equalization in the seventies raised the data rate

of a 3 KHz leased line from 1200 baud to 9600 baud; enhancement by trellis coding in the eighties has produced transmission rates as high as 19.2 Kbaud.

- For satellites, convolutional codes with soft decision decoders have reduced the link budget requirement for digital communication by 6 dB, thus ensuring the feasibility [7] of a new billion dollar industry, namely, VSAT's providing up to 64 kbps service using only 1.2m to 2.4m dishes. The impact of this technology on satellite digital TV broadcasting is just over the horizon. Mobile satellite communication in the heretofore underused Ku-band spectrum is even more impressive. A one-watt transmitter provides both position-location and messaging communications over one transponder of a conventional non-processing satellite [8] for tens of thousands of trucks on North American highways today; the same technology will be available in Europe, and possibly Japan, in the near future.

The second lesson of information theory is just now being learned. Digital compression of speech and video has been a research topic for at least three decades. Bit rate reductions of almost two orders of magnitude are now the norm. High definition television (HDTV) broadcasting is now a major field of research and development at the world's largest electronic manufacturers. Until recently an all-digital approach, which completely separates video-source compression coding from transmission coding, was scoffed at as hopelessly impractical, not because of the feasibility of implementing such compression factors, but because of the difficulty in transmitting 20 to 60 Mbps of digital information, and especially because of the cost of consumer receivers for such.

In 1989, the U.S. Defense Advanced Research Project Agency (DARPA) awarded three development contracts for HDTV signal processing technology, including both image compression and transmission systems. A team headed by Sarnoff Laboratories and a team headed by MIT each received an award, with the third award going to QUALCOMM, Inc. [9] Of these, the first two proposed analog-digital hybrid methodologies, while the third was the only one advocating an all-digital approach. [10]

Since then all known participants have joined the "all digital bandwagon," including General Instruments, NBC Thomson S.A., Philips N.V., Sarnoff, and, more recently, the AT&T-Zenith joint adventure and the MIT-General Instruments "American Television Alliance." [11] Some are concentrating on DBS while others concentrate on terrestrial or cable delivery. All seem to be learning the second lesson.

Multiple Access: A More Subtle Lesson

Which brings us finally to Shannon's third lesson, not yet as widely accepted as the first two. For nearly half a century, military communication specialists, who deal in particularly pernicious and

malignant interference, have intuitively understood that by proper processing at both transmitter and receiver, these could be tamed into behaving like the most benign of interference, thermal (White Gaussian) noise. (It is, in fact, no coincidence that just prior to the research that led to his 1948 publication, Shannon's primary preoccupation was military communication and secrecy.) [12] But personal, wireless and cellular communication also must deal with a variety of interference forms, primarily "the four multiples": multiple-user access, multiple cell-sites, multipath and multiple media.

Through the use of spread spectrum techniques, the detrimental effect of all four can be mitigated, and in some cases (multiple cell-sites and multipath in particular), can even be used to improve communication performance. For example, with wideband spread spectrum modulation, the multiple paths of multipath propagation can be isolated and through diversity combining can be used to advantage. Actually, this is but one of several diversity techniques utilized in a well-designed spread spectrum cellular system. Another is cell-site antenna diversity combining, with a third being the combining of signals from and to two or more cell sites through a technique known as "soft handoff." Equally important unique features of spread spectrum CDMA are interference reduction through "voice activity gating" and "cell-site sectorization gain." What is left, after these mitigating techniques have been exhausted, is the benign additive Gaussian noise, which is itself mitigated by FEC coding.

Most important is the gain in "Reuse Factor," a cellular parameter that is inversely proportional to the number of different frequency assignments necessary to guarantee that neighboring cells are assigned disjoint frequency bands. Reuse factors of 1/19, 1/12, 1/7, 1/4 and 1/3 have been used or proposed for progressively more optimistic designs of FDMA and TDMA systems. [13] CDMA systems employ ubiquitous frequency reuse and thus have a factor of 1. [14] Ubiquitous and universal frequency reuse applies to all users in the assigned spectrum or channel, to all cell sites, and to all media, terrestrial and satellite, both geostationary and low earth orbit. In fact, with CDMA, seamless handover between media becomes possible.

There is, however, one major impediment to universal reuse that must be overcome, particularly in terrestrial systems. Known as the "Near-Far" problem, this refers to the condition in which some users are much closer to the base station than others, thus introducing excessive interference. This has been solved in cellular systems by implementing several levels of rapid, tight, power control. Experimental systems have demonstrated control of power up to 100 dB in fractions of a second, with a total variation in controlled power on the order of + 1.5 dB.

This is not the forum for detailed analyses. Such have been presented by my associates and myself in other venues [15] and publications. [16] To provide some measure of the capacity of CDMA, we quote the current goal of the QUALCOMM CDMA system:

$$\text{CAPACITY (CDMA)} \approx 1 \text{ Bit/Sec/Hz/Cell}$$

This presupposes a Voice Activity factor of 1/2 and Sectorization Gain on the order of 4 to 6

Advances in digital communication in the latter half of this century were guided by the lessons of information theory but fueled by the progress in solid state electronics.

Whichever multiple access technique is employed, the ultimate performance limitation is the system's susceptibility to interference.

dB. (This is by no means a physical limitation; increases of 2 to 4 or even more may well be feasible in second generation systems.)

By comparison, the current equivalent capacity of analog AMPS (assuming the same 10 Kbit/sec voice quality) is

$$\text{CAPACITY (AMPS)} \approx 1/21 \text{ Bit/Sec/Hz/Cell}$$

(Narrow-band AMPS may, at best, double this capacity given that it triples the calls per current channel, but it will undoubtedly require a lower reuse factor.) North American TDMA will provide at best

$$\text{CAPACITY (N.A.—TDMA)} \approx 1/7 \text{ BIT/SEC/HZ/CELL}$$

assuming the current reuse factor 1/7, provided the current C/I -18 dB requirement can be tolerated. European (GSM) TDMA is designed to

$$\text{CAPACITY (GSM-TDMA)} \approx 1/10 \text{ BIT/SEC/HZ/cELL}$$

assuming a reuse factor of 1/4.

What is it about TDMA that prevents it from achieving a higher reuse factor? In TDMA, users in the same cell are kept from interfering (much) with one another by assigning them disjoint frequency and time slots. Once all frequency slots assigned to a given cell have been allocated, the neighboring cells must be assigned another set of frequency slots and so forth, until we reach a distant enough cell that the original frequency can be reused. What keeps this from happening too soon (and therefore with higher reuse factor) is that both the multi-user interference and the multipath interference (that which has not been neutralized by equalization) are seriously detrimental, unlike the benign noise interference characteristic of CDMA.

In short, the principal drawback of TDMA (and FDMA) is that while its proponents have learned the first two lessons of Information Theory (they do employ soft decision decoding and equalization, and they do separate voice compression Vocoding from channel transmission processing and FEC soft-decision decoding), they have missed the third and most important lesson on rendering the interference benign. This is only possible with CDMA and spread spectrum, the logical choice for personal, mobile and wireless digital access.

Economic Considerations

Though capacity and call quality may be the primary concern, other desirable features in personal and mobile communication include: transmitter power requirements of subscriber units; cell-site costs, which are dominated by RF and analog circuitry; transition plan for gradual and profitable conversion from analog cellular and coexistence with existing systems; and security and privacy.

CDMA excels over all other multiple access techniques in providing the best solution for all of the above. Without embarking on a detailed exploration of why and how, suffice it to say again that all are facilitated by the wideband benign interference properties of CDMA of which we have spoken, and that universal frequency reuse avoids the complicated issue of frequency-management plan-

ning when additional cells are introduced.

No assessment is complete without a comparison of implementation complexity and cost. Doubtless, CDMA is conceptually more difficult to understand. But difficulty of concept should not be confused with difficulty of implementation. Were this so, the CD-player would never have become the popular low priced consumer product it is. And this brings me back to my original premise, that advances in digital communication in the latter half of this century were guided by the lessons of information theory but fueled by the spectacular progress in solid state electronics. Specifically, the "smart" and "difficult" algorithms are relegated to the solid state circuitry, where levels of integration and speed double every two years, making yesterday's hopelessly complex and costly implementation into the high volume low cost microchips and ASIC implementations of today. What differentiates CDMA from other multiple access techniques is mostly contained in these chips. At the same time, the conventional analog and RF circuitry in the cell site is greatly reduced because it is shared among more users, with user separation performed at baseband, again in the same powerful chips. The subscriber equipment complexity is also reduced, primarily because of reduced transmitter power requirements.

Let me conclude with a personal observation. I have been privileged to participate in the communication engineering profession and the telecommunication industry over the majority of the period during which we have learned how to apply the key lessons of Shannon's Information Theory. I firmly believe that the final decade of the half century will bring this knowledge to complete fruition in the form of ubiquitous digital communication products and services undreamed of as I entered the field.

References

- [1] K. Kobayashi, *Computers and Communications: A Vision of C & C*, MIT Press, Cambridge, MA, 1986.
- [2] C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, Vol. 27, pp. 379-423 and 623-656, Ju and October, 1948.
- [3] J. A. Heller and I. M. Jacobs, "Viterbi Decoding for Satellite and Space Communication," *IEEE Transactions on Communications Technology*, Vol. COM-19, pp. 835-848, October, 1971.
- [4] G. D. Forney, Jr., "Maximum-Likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference," *IEEE Transactions on Information Theory*, Vol. IT-18, pp. 363-378, May, 1972.
- [5] G. Ungerboeck, "Channel Coding with Multilevel/Phase Signals," *IEEE Transactions on Information Theory*, Vol. IT-28, pp. 55-67, January, 1982.
- [6] R. Wood, "Magnetic Megabits," *IEEE Spectrum*, Vol. 27, pp. 32-38, May, 1990.
- [7] Having previously doubled the communication ranges of space missions such as Voyager and Galileo.
- [8] I. M. Jacobs et al., "The Application of a Novel Two-Way Mobile Satellite Communications and Vehicle Tracking System to the Transportation Industry," *IEEE Transactions on Vehicular Technology*, Vol. 40, pp. 57-63, February, 1991.
- [9] QUALCOMM, Inc., "Technical Proposal for an HDTV Receiver/Processor," submitted to DARPA, Arlington, VA, February, 1989.
- [10] United Press International National Wire Service, "Pentagon Names More HDTV Contractors," October, 26, 1989.
- [11] *New York Times*, "Advanced TV Testing Set Amid Tumult on Technology" pp. C1 and C6, November 15, 1990 and "AT&T and Zenith in Venture Plan to Jointly Build all Digital System for Advanced TV," pp. C1 and C18, December 18, 1990.
- [12] R. Price (Ed.), "A Conversation with Claude Shannon," *IEEE Communications Magazine*, Vol. 22, pp. 123-126, May, 1984.
- [13] W. C-Y Lee, *Mobile Cellular Telecommunications Systems*, McGraw-Hill, New York, 1989.
- [14] This factor is effectively reduced by the increase in interference from neighboring cells, but still remains greater than 3/5. See also Reference [16].
- [15] IEEE GLOBECOM Workshop on CDMA in Satellite and Terrestrial Applications, San Diego, CA, December, 1990.
- [16] K.S. Gilhousen, et al., "On the Capacity of a Cellular CDMA System," *IEEE Transactions on Vehicular Technology*, Vol. VT-40, May, 1991.