# CDMA/HDR: A Bandwidth-Efficient High-Speed Wireless Data Service for Nomadic Users

*Paul Bender, Peter Black, Matthew Grob, Roberto Padovani, Nagabhushana Sindhushayana, and Andrew Viterbi, QUALCOMM, Incorporated*

## ABSTRACT

This article presents an approach to providing very high-data-rate downstream Internet access by nomadic users within the current CDMA physical layer architecture. Means for considerably increasing throughput by optimizing packet data protocols and by other network and coding techniques are presented and supported by simulations and laboratory measurements. The network architecture, based on Internet protocols adapted to the mobile environment, is described, followed by a brief discussion of economic considerations in comparison to cable and DSL services.

## INTRODUCTION

The rapid growth and nearly universal coverage of industrialized nations and regions by digital wireless telephony gives rise to an increasing demand for data services as well. While current offerings are for data rates equivalent to those provided by wireline modems a decade or more ago, the gap is closing. Standards are already approved and chip sets are available for providing data rates above 64 kb/s within this calendar year (and century). Just beyond this horizon, however, service providers are already planning for wireless data rates above 2 Mb/s, approaching those of wireline, digital subscriber line (DSL), and cable. Whether such a wireless service can be made technically and economically competitive with wireline and cable is not the main issue, although we shall address this briefly in the last section. What will drive such a service is the demand for rapid low-latency availability of Internet access to nomadic users. In the next section we describe the characteristics and perceived needs of this user class. We then proceed to explore the characteristics of data requirements for speed and latency, after which we present a technical system solution tailored to these requirements and the characteristics of a specific implementation as an evolution of existing CDMA base station and subscriber terminal architectures. In the final section we briefly discuss the economics of such a system deployment.

## CHARACTERISTICS OF NOMADIC USER DATA DEMAND

In business and the professions, the individual is often absent from her or his normal workplace. To continue to be productive on the road, both in transit and at business or professional meetings, connectivity to data at one's principal workplace and more broadly to other databases accessible through the Web is essential. Generally, members of this nomadic user class demand the same data service normally available in their home base. Often cited examples of the nature of such services are e-mail retrieval, Web browsing, ordering airline tickets, hotel reservations, obtaining stock quotes, and report retrieval, in such locations as airport lounges, hotel rooms, and meeting places, in each case without recourse to the limited or interface unfriendly facilities available in such places. In fact, one need not necessarily look beyond corporate boundaries; professional employees often spend nearly as much time in company conference rooms as in their own offices, and rarely are such rooms equipped with the number of ports needed to connect the majority of participants' laptops.[1]

The nature of such data traffic is decidedly asymmetric. A much higher forward (or downlink) rate is required from the access point (base station) than that generated by the access terminal (user terminal) in the reverse (uplink) direction. Furthermore, just as does the fixed user, the nomadic user expects a response to her or his request which does not suffer from excessive latency. Our goal is to satisfy these needs with an evolutional approach which minimizes the time and cost for providing such capabilities in existing cellular infrastructure and with terminals that differ only at digital baseband from existing cellular and personal communications systems (PCS) handsets.

[1] *The obvious alternative of private campus wireless systems will be discussed in the last section.*

## The Technical Challenge

Little information has been gathered on the exigencies of nomadic users and the networks to serve them, simply because, with the exception of a very small percentage of low-speed data service, such networks have not existed. On the other hand, with digital cellular networks in place for nearly a decade and with large numbers of mobile users served for several years, a great deal is known about the characteristics of digital wireless networks and their mobile users. A major step in the perfection of digital cellular technology was the development and standardization of code-division multiple access (CDMA) wireless systems and their adoption by the majority of North American, Korean, and Japanese carriers and manufacturers. Having proven its superiority to other access techniques, it is now being imitated (sometimes in a modified form) by most of the carriers and manufacturers who were the initial holdouts and skeptics of its viability.

CDMA was designed for efficient reverse (uplink) and forward (downlink) operations. It was initially widely believed that the reverse direction in which multiple users access each base station, hence representing multiple sources of interference with one another, would be the capacity limiting (or bottleneck) direction. This assumption turned out to be incorrect; the forward (downlink) was the initial bottleneck for three principal reasons:

- Interference on the reverse link enjoys the advantage of the law of large numbers, whereby the cumulative interference from multiple low-power transmitters tends to be statistically stable. The forward link, on the other hand, suffers interference from a small number of other high-power base stations. This becomes particularly serious at the vertices of the (imaginary) cellular hexagon where the transmitting base station and two other interfering base stations are equidistant from the intended user. This situation is relieved by soft handoff, where two or more base stations transmit to the user simultaneously.
- But soft handoff, while greatly diminishing interference, which itself increases capacity, still overall diminishes the forward link[2] capacity because an additional CDMA carrier must be assigned in the newly added base station. Depending on the region of (or criterion for) soft handoff, this can cause greater or lesser reduction.
- While on the reverse link, fast and accurate power control of multiple users is evidently critical to operation and capacity realization, it was initially felt in producing the first CDMA standard, cdmaOne (IS-95-A [1]), that forward link power control could be much slower. This turned out to reduce forward link capacity.

The second and third limiting causes have been eliminated or considerably diminished in the evolutionary revisions of cdmaOne (IS-95-B and CDMA2000). Fast power control is now implemented in the forward link, and the region of (criterion for) soft handoff has been diminished. The first cause (sometimes called the "law of small numbers"), remains, however. These improvements have brought the forward (downlink) capacity to parity with the reverse (uplink) capacity. But for high-speed data, such as downloading from the Internet, this is *not enough*. The downlink demand is likely to be several times greater than the uplink. The rest of this article deals with new approaches which will further increase the downlink capacity by a factor of three to four for data applications only.[3]

## The Technical Approach to High-Speed Data

Most data applications differ fundamentally from speech requirements in two respects already noted, traffic asymmetry and tolerance to latency. Two-way conversational speech requires strict adherence to symmetry; also, latencies above 100 ms (which corresponds to about 1 kb of data for most speech vocoders) are intolerable. For high-speed data downlinked at 1 Mb/s, for example, 100 ms represents 100 kb or 12.5 kbytes; furthermore, latencies of 10 s are hardly noticeable, and this corresponds to a record of 1.25 Mbytes. Thus, smoothing over a variety of conditions, which is always advantageous for capacity, is easily accomplished.

All communication systems, wired as well as wireless, are greatly improved by a combination of techniques based on three principles:

- Channel measurement
- Channel control
- Interference suppression and mitigation

Our approach employs all three. First, on the basis of the received common pilot from each access point (or base station), each access terminal (subscriber terminal) can measure the received signal-to-noise-plus-interference ratio (SNR). The data rate which can be supported to each user is proportional to its received SNR. This may change continuously, especially for mobile users. Thus, over each user's reverse (uplink) channel, the SNR or equivalently the supportable data rate value is transmitted to the base station. In fact, since typically two or more base stations may be simultaneously tracked, the user indicates the highest among its received SNRs and the identity of the base station from which it is receiving it, and this may need to be repeated frequently (possibly every slot[4]). In this way the downlink channel is controlled as well as measured. Furthermore, by selecting only the best base station, in terms of SNR, to transmit to the user, interference to users of other base stations is reduced. Additionally, since data can tolerate considerably more delay than voice, error-correcting coding techniques which involve greater delay, specifically turbo codes, can be employed which will operate well at lower $E_b/N_0$, and hence lower SNR and higher interference levels.

Next, we show how unequal latency, for users of disparate SNR levels, can be used to increase throughput. Suppose we can separate users into $N$ classes according to their SNR levels, and corresponding instantaneous rate levels supportable. Thus, user class $n$ can receive slots at rate $R_n$ b/s, where $n = 1.2..N$, and suppose the relative frequency of user packets of class $n$ is $P_n$.

*All communication systems, wired as well as wireless, are greatly improved by a combination of techniques based on three principles: channel measurement, channel control, and interference suppression and mitigation Our approach employs all three.*

[2] *Although not the capacity of the reverse link, which soft handoff actually increases.*

[3] *Clearly voice is fundamentally a symmetric service, with stricter latency requirements, as we shall note below.*

[4] *In speech-oriented CDMA, voice frames are 20 ms long. In the next section, we shall establish corresponding lengths for data, which will be called slots. Multiplying* $R_{av}$ *of (1) by slots per second yields throughput in bits per second.*

| Data rate (kb/s) | Packet length (bytes) | FEC rate (b/sym) | Modulation |
|---|---|---|---|
| 38.4 | 128 | 1/4 | QPSK |
| 76.8 | 128 | 1/4 | QPSK |
| 102.6 | 128 | 1/4 | QPSK |
| 153.6 | 128 | 1/4 | QPSK |
| 204.8 | 128 | 1/4 | QPSK |
| 307.2 | 128 | 1/4 | QPSK |
| 614.4 | 128 | 1/4 | QPSK |
| 921.6 | 192 | 3/8 | QPSK |
| 1228.8 | 256 | 1/2 | QPSK |
| 1843.2 | 384 | 1/2 | 8PSK |
| 2457.6 | 512 | 1/2 | 16QAM |

**Table 1.** *Various data rates.*

Suppose slots are assigned one at a time successively to each user class. Then the average rate, which we define as throughput, is

$$R_{av} = \sum_{n=1}^{N} P_n R_n \text{ b/s.} \tag{1}$$

This, of course, means that lower-data-rate (and SNR) users will have proportionately higher latency. For if $B$ bits are to be transmitted altogether for each class, the number of slots (and hence time) required for user class $n$ will be $B/R_n$, and hence the latency $L_n$ is inversely proportional to $R_n$.

Suppose, on the other hand, that we require all users to have essentially the same latency[5] irrespective of the $R_n$ they can support. Then as each user class is served, it will be allocated a number of slots inversely proportional to its rate. Let $F_n$ be the number of slots allocated to class $n$, where $F_n = k/R_n$, $k$ being a constant. In this case, the average rate or throughput is

$$R'_{av} = \frac{\sum_{n=1}^{N} P_n R_n F_n}{\sum_{n=1}^{N} P_n F_n} = \frac{1}{\sum_{n=1}^{N} P_n / R_n} \text{ b/s.} \tag{2}$$

In this case, however, the latency of all user classes will be the same (assuming the total number of bits $K$ to be large and thus ignoring edge effects).

To assess the cost in throughput for equalizing latency, consider the extreme case of only two user classes, each equally probable ($P_1 = P_2 = 1/2$) but capable of supporting very disparate rates $R_1 = 16$ kb/s, $R_2 = 64R_1 = 1{,}024$ kb/s. Then in the first case, $R_{av} = 520$ kb/s. but $L_1/L_2 = 64$. In the second case, $L_1 = L_2$, but $R'_{av} = 31.51$ kb/s.

To see that there is a more rational allocation which is less "unfair" than a latency ratio of 64, and still achieves a better throughput than $R'_{av}$, consider a compromise which guarantees that the highest latency is no more than, for example, 8 times the lowest latency. Then in the second case, we would assign 8 slots to class 1 for every slot assigned to class 2. The result would be $L_1/L_2 = 8$ as required and

$$R_{av}'' = (8FP_1R_1 + FP_2R_2)/(FP_1 + FP_2) = 128 \text{ kb/s.}$$

For the general case of $N$ classes and latency ratio $L_{max}/L_{min}$, it can be shown that the maximum achievable throughput, denoted by $C$, is

$$C = \frac{\sum_{n=1}^{n_o} P_n + \sum_{n=n_o+1}^{N} P_n (L_{min}/L_{max})}{\sum_{n=1}^{n_o} P_n / R_n + \sum_{n=n_o+1}^{N} (P_n/R_n)(L_{min}/L_{max})} \text{ b/s,} \tag{3}$$

where $R_1 < R_2 \ldots R_N$ and $n_o$ is such that $R_n \le C$ for all $n \le n_o$, while $R_n > C$ for all $n > n_o$.

Surprisingly, with this maximizing strategy, each user's latency is either $L_{max}$ (for those for which $R_n < C$) or $L_{min}$ (for those for which $R_n > C$). To determine the maximum throughput it is necessary to have a histogram of the achievable rates for users of the wireless network in question. This will be discussed in the next section. Also, as we shall find there, practical numerology considerations may require us to deviate from this strict bimodal latency allocation, although the ratio $L_{max}/L_{min}$ will remain as the principal constraint.

## Implementation of High-Data-Rate Code-Division Multiple Access

In the last section we discussed the key factors and parameters of a wireless system designed to optimize the transport of packet data. In the following we will describe such a system design, beginning first with a description of the air interface, to continue in the next section with a description of the network architecture. The design leverages in many ways the lessons learned from the development and operation of CDMA IS-95 networks, but makes no compromises in optimizing the air interface for data services. Furthermore, a compelling economic argument can be made for a design that can reuse large portions (to be exact, all but the baseband signal processing elements) of components and designs already implemented in IS-95 products, both in the access terminals and access points (APs).

Due to the highly asymmetric nature of the service offered, we will focus most of our attention on the downlink. In the IS-95 downlink, a multitude of low-data-rate channels are multiplexed together (with transmissions made orthogonal in the code domain) and share the available base station transmitted power with some form of power control. This is an optimal choice for many low-rate channels sharing a common bandwidth. The situation becomes less optimal when a low number of high-rate users share the channel. The inefficiencies increase further when the same bandwidth is shared between low-rate voice and high-rate data users, since their requirements are vastly different, as discussed previously. It should be noted that

[5] *This is the case for voice. The only difference is that in speech, transmitter power levels are controlled to equalize received power, while here time, in terms of frames, is controlled to equalize energies.*

increasing the bandwidth available for transmission cannot help in this regard if the data rate of the users is increased proportionally as well.

Therefore, a first fundamental design choice is to separate the services, that is, low-rate data (voice being the primary service in this category) from high-rate data services, by using possibly adjacent but nonoverlapping spectrum allocations. To summarize, a better system is one that uses an IS-95 or cdma2000-1X RF carrier to carry voice and a separate high-data-rate (HDR) RF carrier to deliver high-rate packet bursts.

With a dedicated RF carrier, the HDR downlink takes on a different form than that of the IS-95 designs. As shown in Fig. 1, the downlink packet transmissions are time multiplexed and transmitted at the full power available to the AP, but with data rates and slot lengths that vary according to the user channel conditions. Furthermore, when users' queues are empty, the only transmissions from the AP are those of short pilot bursts and periodic transmissions of control information, effectively eliminating interference from idling sectors.

The pilot bursts provide the access terminals with means to accurately and rapidly estimate the channel conditions. Among other parameters, the access terminal estimates the received $E_c/N_t$ of all resolvable multipath components and forms a prediction of the effective received[6] SNR. The value of the SNR is then mapped to a value representing the maximum data rate such a SNR can support for a given level of error performance. This channel state information, in the form of a data rate request, is then fed back to the AP via the reverse link data rate request channel (DRC) and updated as fast as every 1.67 ms, as shown in Fig. 2. The reverse link data request is a 4-bit value that maps the predicted SNR into one of the data rate modes of Table 1. In addition, the access terminal requests transmission from only one sector (that with the highest received SNR) among those comprising the *active set*. Here the definition of active set is identical to that for IS-95 systems, but unlike IS-95, only one sector transmits to any specific access terminal at any given time.

The main coding and modulation parameters are summarized in Table 1.

The forward error correcting (FEC) scheme employs serial concatenated coding and iterative decoding, with puncturing for some of the higher code rates [2].

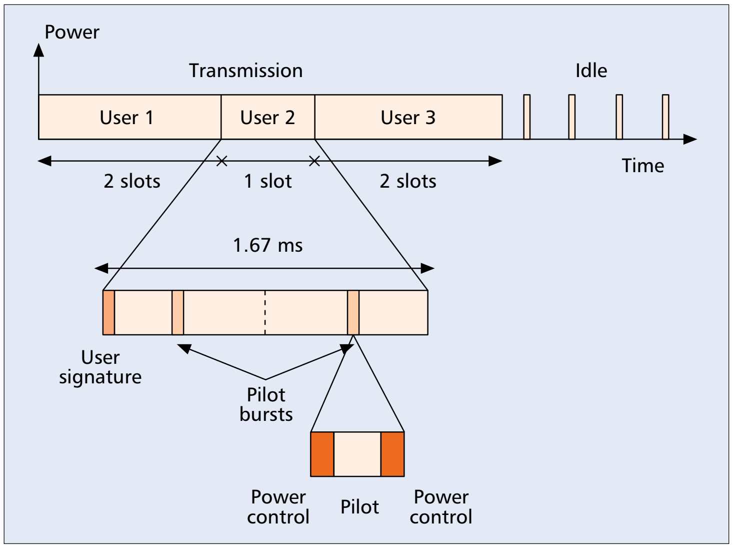


**Figure 1.** *An access point transmission diagram.*

Following the encoder, these traditional signal processing steps are applied: symbol repetition is performed on the lower-data-rate modes; scrambling, channel interleaving, and the appropriate modulation is applied to obtain a constant modulation rate of 1.2288 MHz for all modes. The in-phase and quadrature channels are then each demultiplexed into 16 streams, each at 76.8 kHz, and 16-ary orthogonal covers are applied to each stream. The resulting signal, obtained by adding the 16 data streams, is then spread by quadrature pseudonoise (PN) sequences, bandlimited and upconverted. The resulting RF signal has the same characteristics as an IS-95 signal, thus allowing the reuse of all analog and RF designs developed for IS-95 base stations, including the power amplifiers, and the receiver designs for subscriber terminals.

Table 2 summarizes the SNR required to achieve a 1 percent packet error rate (PER).

Note that at the lower rates this corresponds to $E_b/N_0 \approx 2.5$ dB, a result of using iterative decoding techniques on serial concatenated codes, while for the two highest rates, $E_b/N_0$ increases considerably because 8-phase shift keying (PSK) modulation and 16-quadrature amplitude modulation (QAM) are employed. These were obtained both by bit-exact simulation and

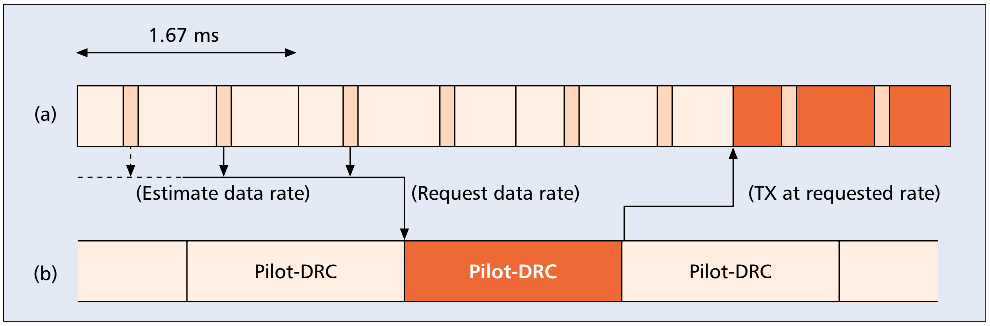


**Figure 2.** *A channel estimation and data request channel timing diagram: a) access terminal receive; b) access terminal transmit.*

[6] $E_c$ *represents the received signal energy density and* $N_t$ *represents the total nonorthogonal single sided noise density.* $N_t$ *comprises intercell interference, thermal noise, and possibly nonorthogonal intracell interference.*

corroborated by laboratory measurements with a complete RF link.

At this point we are able to estimate the maximum achievable throughput per sector as discussed in the previous section. Figure 3a shows a graph of the cumulative distribution function of the SNR for a typical embedded sector of a large three-sector network deployed with a frequency reuse of one. In particular, the SNR values are those of the best serving sector and representative of a uniform distribution of users across the coverage area. From the results of Fig. 3a and the knowledge of the SNR required to support a given data rate (Table 1), it is straightforward to derive the histogram of data rates achievable in such an embedded sector. The result is shown in Fig. 3b where the SNRs used in the calculation are those of Table 1 with an additional 2 dB of margin to account for various losses. Finally, Fig. 3c shows the realized throughput per sector per 1.25 MHz of bandwidth versus the parameter $L_{max}/L_{min}$. Note that the throughput is doubled for a latency ratio $L_{max}/L_{min} = 8$.

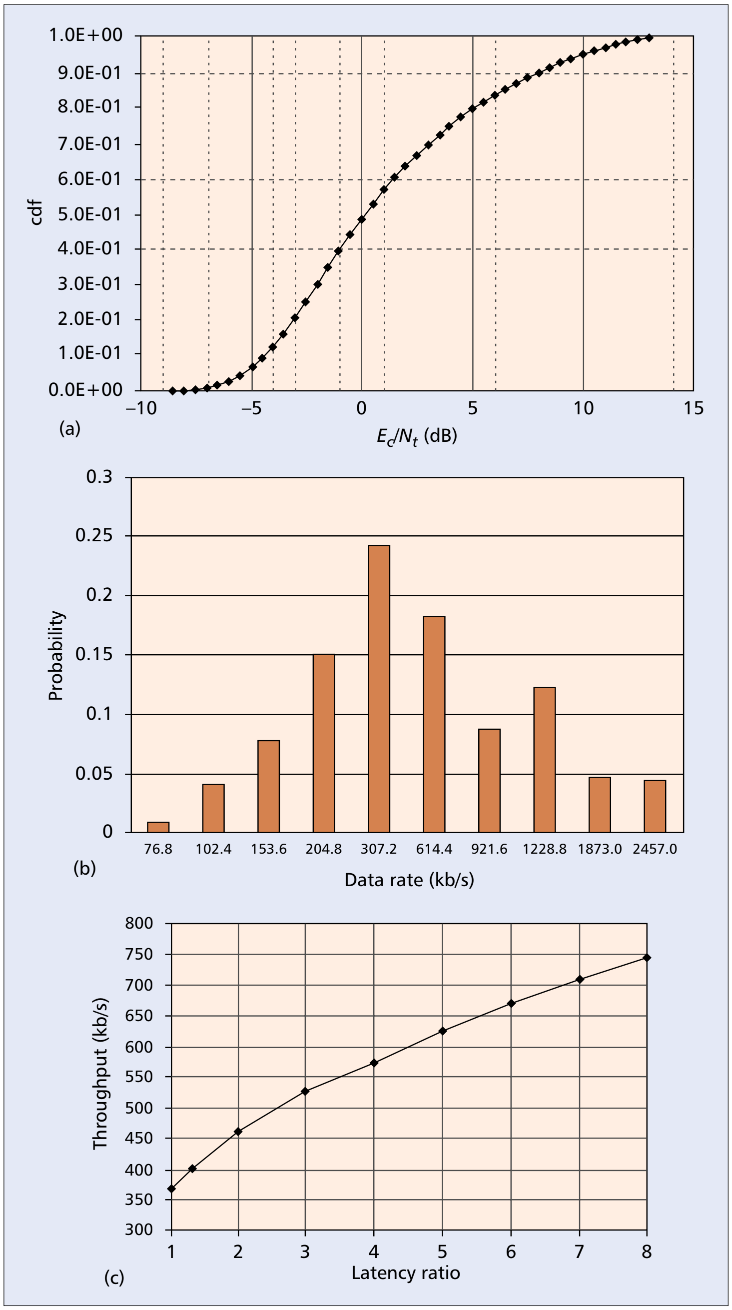


**Figure 3.** *a)* $E_c/N_t$ *distribution for a typical embedded sector in a three-sector network with universal frequency reuse in each cell; b) data rate histogram; c) sector throughput vs. latency ratio* $L_{max}/L_{min}$.

## Network Architecture

Since the radio link has been designed to provide efficient access to packet data networks, it is natural to turn to the most ubiquitous packet data network — the Internet — when selecting the network architecture. Adopting Internet protocols in the communication between the access terminal and the access network allows users to access the widest variety of information and services, including e-mail, private intranets, and the World Wide Web. Furthermore, the selection of Internet protocols in the design of the access network allows the access network equipment to take advantage of the ever decreasing costs and increasing performance of Internet equipment.

First, we examine the communication link between the access terminal and the access network. Figure 4a shows the protocol stack used in such a link.

In order to carry traffic between the user and the network, we need to select a network-layer protocol. We chose the Internet Protocol (IP) [3] because it is the network-layer protocol of the Internet. The Internet carries its network-layer protocol over a variety of transports. For example, asynchronous transfer mode (ATM) often carries Internet traffic on the Internet backbone, Ethernet often carries Internet traffic on local area networks (LANs), and the Point-to-Point Protocol (PPP) [4] often carries Internet traffic over dialup connections. We chose PPP for the following reasons. First, PPP is widely supported. Moreover, PPP allows the transport of a variety of network-layer protocols, supports methods for

| Data rate (kb/s) | $E_c/N_t$ (dB) |
|---|---|
| 38.4 | –12.5 |
| 76.8 | –9.5 |
| 102.6 | –8.5 |
| 153.6 | –6.5 |
| 204.8 | –5.7 |
| 307.2 | –4.0 |
| 614.4 | –1.0 |
| 921.6 | 1.3 |
| 1228.8 | 3.0 |
| 1843.2 | 7.2 |
| 2457.6 | 9.5 |

**Table 2.** *SNR for a 1 percent packet error rate.*

# Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

## Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

## Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

## Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

## API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

**LAW FIRMS**

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

**FINANCIAL INSTITUTIONS**

Litigation and bankruptcy checks for companies and debtors.

**E-DISCOVERY AND LEGAL VENDORS**

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | sales@docketalarm.com | 1-866-77-FASTCASE