A New ALERT Protocol

Feasibility Study of a New Air Interface
and
Physical Layer Packet Definition
for
The ALERT User Community

R. Chris Roark
Blue Water Design LLC
Niwot, Colorado

and

Donald G. Van Wie
DIAD Incorporated
Longmont, Colorado

February 2003

Blue Water Design, LLC
Technology & Business Development

DIAD

This page intentionally left blank.

Table of Contents

This page intentionally left blank.

Executive Summary

This study was undertaken by Blue Water Design, LLC with support from DIAD Inc. under contract to the ALERT User Group to investigate whether it is feasible to increase the data rate of the ALERT protocol and to survey what solutions exist in the current data radio communications market.

The general requirements of a new ALERT protocol – to support more sensors, send higher resolution data, detect and correct errors – require sending more information. The only option for sending more information without decreasing channel capacity is to increase the data rate.

The air interface problem

The feasibility question – Can we significantly increase the data rate of ALERT messages without degrading the current performance? – hinges on two radio frequency propagation issues. The first is whether we can overcome the loss of bit energy that inevitably follows from increasing the data rate. The second major issue is whether we can manage the effects of multipath interference at higher bit rates.

The current ALERT message is transmitted at 300 bits per second (bps). Each bit can be thought of as a pulse of RF energy that lasts 3.33 milliseconds (msec). It is this energy that is detected at the receiver, and that must be identifiable relative to the background radio noise. If the same transmitter is used to send a message at 2400 bps, the bit duration falls to less than 0.42 msec, and the energy per bit falls by a factor of 8. If we assume the same environmental noise level and that the transmitter power cannot be increased, we must have circuitry that can reliably detect the message when the signal-to-noise ratio (S/N) is 8 times weaker.

Multipath, as the name suggests, is the term applied to the effects at the receiver when a transmitted signal arrives via more than one path. The shortest path is the line of sight between transmitter and receiver, but the radiated signal may be refracted by the terrain or it may be reflected off buildings, mountains or valley walls to eventually arrive at the same place. Signals that travel different length paths arrive at different times. If the delay is long enough, the energy from adjacent bits may overlap, perturbing the ability to distinguish bit values. Each doubling of the bit rate halves the difference in path length that will produce the same degree of inter-symbol interference. The existing ALERT has been relatively immune to multipath problems because its symbol period of 3.3 milliseconds is very long in relation to the travel time differences of alternate paths. As the symbol period decreases with increasing bit rate, multipath must be accounted for in assuring communication reliability

Summary of Findings

The authors have examined these issues in some detail, and conclude that moving to 3600 baud is feasible and practical, provided more sophisticated modulation and data coding technologies are used. A transmission rate of 4800 baud appears possible but its feasibility cannot be assured. The impediment to 4800 baud would be the inter-symbol interference caused by multipath and we recommend field trials to confirm the suitability of the higher rate.
With regard to the first problem, the loss in energy per bit is 10.8 dB (for 3600 bps) to 12 dB (for 4800 bps). This paper examines what methods can be used to recover this loss through gains in power efficiency, while continuing to use the inexpensive, off the shelf transceivers of the current ALERT world. The gains are made through a combination of three methods:

- First, the current ALERT Audio Frequency Shift Keying (AFSK) developed for wire-line transmission would be replaced by the carrier FSK currently practiced in the radio telemetry domain. Presently, two audio tones are used to signify the binary values in a message; these audio tones are fed to an FM modulator that varies the FM frequency in relation to the changing audio voltage. At the receiver, the signal is demodulated twice, first from a varying frequency (FM) signal to an analog audio signal in the transceiver, then from the audio signal to a binary logic voltage in the ALERT demodulator (decoder). Current technology would use the binary logic values to drive the carrier between two discrete frequencies (FSK), which gives the design engineer control over factors which optimize the efficiency of data throughput.

- The second method involves filtering the digital input signal to eliminate instantaneous frequency shifts between bits. Rapid shifts cause the RF energy to be scattered over more of the spectrum in the form of side lobes to the main energy lobe. Gaussian, Raised Cosine or other forms of filtering, along with judicious choice of shift frequencies, reduce the amount of side lobe energy and focus it in the central portion of bandwidth. We estimate that the combination of the first two methods, which are changes to the modulation scheme, will result in efficiency improvements that range from 5 to 8 dB over the current ALERT techniques

- Another 6 to 7 dB of power efficiency can be gained by how the message is encoded. Using convolutional and Reed-Solomon forward error correction codes, combined with bit interleaving, allows the message to be reconstituted at a much lower signal-to-noise ratio. These techniques are relatively simple to implement at the encoding (sensor) end. Decoding is computationally intensive, possibly limiting it to repeaters and base station decoders where the cost and power consumption of higher performance processing is tolerable. Methods for implementing the message coding are available as open-source C program code and can be adapted to one or more microprocessors selected for use in the ALERT domain.

Our theoretical and analytical conclusions are that power efficiency can be improved by the 10 to 12 dB needed to offset the increased bit rate. This would maintain the same bit error rate that characterizes the existing ALERT protocol

The detrimental effects of multipath increase with increasing data rate and may be what ultimately constrains the ALERT bit rate. The significance of multipath varies from site to site and may not be a significant factor in the majority of ALERT installations regardless of data rate. The position of this study is to make a feasibility recommendation based on a principle of no degradation to existing functionality. Based on the limited published data available, and pending field studies in ALERT environments, we believe that a 3600 to 4800 bps transmission rate will be viable.

The effect of multipath is to introduce bit errors similar to white noise. The forward error correction proposed to offset the reduced energy per bit will work to overcome the effects of multipath. Additionally, we propose considering a 4-level FSK coding scheme to permit a 4800 bps throughput at an over the air rate of 2400 baud.

The feasibility study predicts that ALERT data can be moved at 12 to 16 times the present rate without increasing the bit error rate. An additional benefit that we could not quantify is that virtually all errors can be detected. Corrupted reports that cannot be corrected can be removed from the data stream. In the ALERT monitoring world, the user is much more severely impacted by corrupted data than by missing data. The latter is expected from time to time given the random event-based broadcast techniques used in ALERT, while the former can cause false alarms and loss of confidence in system integrity

The modulation and coding techniques proposed here are not leading edge. They have been deployed extensively in a variety of packet data networks, including cellular telephony, space communications, paging, trunked radio, and other applications. Devices have been designed and built in the amateur radio domain that interface to off-the-shelf transceivers and communicate at rates up to 9600 bps in 12.5 kHz land mobile channels.

One hope at the outset of this effort was that one or more of the commercial vendors of this technology could be induced to make their methodologies an open ALERT standard. However, manufacturers have shown no interest in this approach. Failing that, we hoped we could find and adopt an existing open protocol. An extensive review of the existing protocols turned up nothing worth adopting wholesale as the new ALERT standard. Although many of the protocols come close to ALERT's needs, they are not suitable because they are optimized for voice or large data streams, or are connection oriented, or carry legacy features from their original applications.

There is a good deal of convergence in the design of wireless protocols, and these principles would be applied in the new ALERT protocol packet structure. The authors propose a skeleton format that is a fixed-length packet that can specify an optional, variable size trailing data block. The packet includes a protocol identifier that signals the structure used in the data block, so the ALERT protocol could become a family of supported formats. The ALERT community can prescribe the specific data elements supported, their sizes, structures and uses as the effort goes forward.

The authors are confident that a new ALERT protocol operating at least 3600 bps is practical and feasible. The next step is to produce a set of working prototypes for bench testing, comparing performance at 3600 and 4800 baud against current ALERT systems. The outcome of bench testing will be a suite of technologies that are ready for testing under typical and extreme ALERT field conditions.

This page intentionally left blank.

Introduction

We undertook this study under a contract with the ALERT Users Group (AUG) and with the support of the Southwestern Association of ALERT Systems, (SAAS). The purpose of this effort is to establish the technical and economic practicality of making significant changes to the ALERT protocol. The users groups funded the study in order to assure that the results of this effort are available throughout the ALERT community and to all vendors serving it.

ALERT is an acronym for Automated Local Evaluation in Real Time. It was developed and introduced by the National Weather Service in the early 1970's to provide rainfall data in real time to both local managers and NWS officials. Although loosely applied to a variety of technologies and practices, in this study ALERT refers specifically to a 300-baud, Audio Frequency Shift Keying over FM, four-byte message protocol for real-time hydrologic data transmission using line of sight radio.

ALERT has been in use, virtually without modification, for nearly 30 years. Over that time its use has expanded from rainfall monitoring to full suites of weather sensors, as well as monitoring water level, soil and fuel moisture, pavement conditions, water quality and other real-time data collection. It is even used in control applications, such as operating warning flashers and roadway gates.

The number of sensors in use now greatly exceeds the ID numbering capacity of the ALERT protocol, forcing ID reuse and complicating the regional integration of data. Users want to integrate intelligent sensors that have high-resolution outputs in engineering units, which challenges the limited capacity of the ALERT data field. The inability to detect errors puts ALERT at a severe disadvantage in the critical applications where its real-time capabilities are most useful.

There are many reasons for contemplating changes to the ALERT protocol. It antedates virtually all of the signaling systems in use today and, while it represented a remarkable advance in hydrologic telemetry in the 1970s, it now constrains its users with low throughput, limited message size, lack of error detection or correction, and limited flexibility.

Virtually all modern radio data systems, be they cellular telephony, supervisory control and data acquisition (SCADA), space communications, or paging, operate at much higher throughputs with significantly reduced error rates. The FCC has responded to the improving capabilities of modern telemetry systems by "narrowbanding", that is, decreasing the allowable bandwidth per channel in order to create more channels and improve spectrum utilization.

The goals of a new ALERT protocol are to:

- Increase the effective data rate in order to increase channel capacity;
- Permit the incorporation of more information into an ALERT message;
- Introduce more flexibility into the structure of an ALERT message for wider applicability of its use;
- Incorporate robust error detection and correction;
- Have a standard that resides in the public domain;

- Use an air interface that can be implemented with generally available components by multiple vendors.

A simple demonstration was performed in 2000 using proprietary radio and modem components. The results strongly suggest that eight-fold to sixteen-fold increases in data rate are possible. The present study encompasses the physical issues that limit the transfer of data by radio telemetry. The next step is to select the most promising methodologies that emerged from this study to build prototypes that can carry a new ALERT protocol to the field trial stage.

The Air Interface – Physical Layer Packet Definition Constraints

If a new ALERT protocol is to be successful, it must meet a number of criteria, including regulatory requirements, cost-benefit profile, and the operational context of the ALERT community. The following are constraints that we believe any new protocol must meet:

- Maintain compatibility with current FCC Land Mobile channel requirements. The new scheme shall not increase effective isotropically radiated power (EIRP) and must operate within 12.5 kHz channel bandwidths.
- Use an ALERT modem that is independent of the RF transceiver so the vendor or user can select among off-the-shelf transceivers. The new modem should have a receive interface at the FM (Frequency Modulated) transceiver's discriminator output, and transmit via the transceiver's audio input.
- Not reduce the probability of error-free ALERT message reception in severe radio terrain environments using UHF and VHF channels.
- Keep increases in ALERT transmitter modem costs and complexity to a minimum.
- Maintain at least the same level of reliability as the current ALERT systems.
- Hold any increase in power consumption at the ALERT transmitter to insignificant levels. At receiver sites (repeater and base station) a power consumption increase may need to be tolerated, but should be minimized.
- Maintain or improve the simplicity of installation and serviceability of the current ALERT networks.
- Enable an incremental implementation of the new ALERT gages; allow the operation of a mix of new and old on a single channel.
- Retain the connectionless, broadcast, self-initiated transmission capabilities of the existing ALERT architecture.

Survey of Current VHF/UHF Integrated Radio Modems

The authors surveyed the current VHF/UHF data radio market to determine the feasibility of using an available integrated radio modem. A representative sampling of the integrated radio modems available and their specifications is compiled in Appendix A.

Most available products incorporate some form of packet protocol. Some acknowledged being a modification of a standard, such as Simrex's MX.25, which is a revision of the AX.25 Amateur Radio protocol that in turn is based on X.25. Teledesign's AIRNET is a Carrier Sense Multiple Access (CSMA) slotted Aloha protocol that appears to draw upon AX.25 concepts, but is significantly different. The vendors' protocols are proprietary, and there is no compatibility among them at the physical packet structure level.

The only inter-vendor compatibility found is at the modulation level, using Gaussian Minimum Shift Keying (GMSK) at a standard rate of 4800 or 9600 bps. This is due to 1) GMSK's proven bandwidth efficiency and 2) availability of implementation technology since the GSM cellular system has standardized on GMSK. No vendor suggested or claimed compatibility with any other vendor's product in any discussion or in any specification sheet. When the authors discussed with vendors the possibility of the ALERT Users Group adopting their protocol and making it an open standard, their universal response was that their proprietary protocol was a "differentiator". Specifically, neither Data Radio nor Simrex would entertain a discussion of opening their protocol.

A representative sampling of prices was obtained for integrated 9600 bps or higher modems with a 2 W UHF radio for potential purchase quantities of 10. Unfortunately the quantities required for a price discount from all the vendors sampled is at either 25 units or 100 units, so no discounts applied.[1] The price ranged from a low of $700 per unit to a high of $1,295 per unit.

We have concluded that a commercially integrated high speed modem and radio is not a viable candidate for the new ALERT protocol, for two reasons:

1. No vendor is willing to move its proprietary protocol into the public domain;

2. There is no interoperability among hardware from different manufacturers.

Nevertheless, the specifications for these integrated units illustrate the current commercial state of the art. Most are capable of 9600 bps in either a 12.5 or 25 kHz channel, and many provide 19,200 bps in a 25 kHz channel. Most also are selectable for 4800 bps. Direct comparisons are difficult because there is no normalized bit error rate (BER) specification. For example, Simrex's Synthesized Netlink Radio Data System (SNRDS) appears to achieve a 6 dB superior BER compared to Data Radio's T96R, with both operating at 9600 bps in a 25 kHz channel. Simrex uses Gaussian MSK, while Data Radio's T96R uses Differential Raised Cosine MSK (DRCMSK). One probable explanation is Simrex's use of a channel code, a Hamming (12,8) code, to correct errors in the transmission.[2]

All data radio-modems in Appendix A use some form of constant envelope Frequency Modulation (FM), either digital Frequency Shift Keying (FSK) or Audio Frequency Shift Keying-FM modulation. No phase shift keying (PSK) modulation or Amplitude Modulation (AM) is used. All higher bit rates use either Gaussian Minimum Shift Keying (MSK), Gaussian Frequency Shift Keying (GFSK), or 4-Level GFSK.

Existing Commercial Communication Protocols

Even though an integrated transceiver and modem is not a viable solution, it may still be possible to take advantage of an existing commercial or standard protocol's Physical Layer Packet and/or modulation technology. A protocol which is in wide use is likely to be implemented in an

[1] One vendor discounted a list price of $1,595 to $1,000 when purchase information was requested. It is unknown whether this was a one-time discount or could be applied generally.
[2] Also note that the (12,8) code is a 50% overhead code, so the SNRDS's effective data throughput is only 6,400 bps.

Application Specific Standard Product integrated circuit (ASSP). Adopting the new ALERT protocol around a high volume ASSP-implemented protocol would simplify development and could lower the new ALERT modem cost.

Suitability for adoption as the new ALERT protocol is dependent on the extent to which it incorporates the following key features:

- Simple multicast short message capability, with the flexibility for future connection-oriented network applications;
- Designed for audio input and limiter discriminator output interfaces to FM VHF/UHF land mobile transceivers;
- Open licensing;
- Forward error correction (FEC);
- Robust synchronization, assuring reliable packet detection in low signal to noise (S/N) environments;
- Variable packet size with fine granularity.

.
There is an extensive, diverse array of public and proprietary wireless packet protocols. The most common are for cellular, paging and private packet systems. Others include telemetry protocols, trunked radio protocols and cordless telephone protocols.[3] Only a handful of these are candidates for detailed evaluation. None of the cellular based protocols such as Cellular Digital Packet Data (CDPD), Digital PCS Short Message System (SMS) and General Packet Radio Service (GPRS) are viable since they all rely on time-slotted or spread spectrum synchronization mechanisms inherent in the two-way cellular infrastructure.

Some aspects of Trunked radio protocols are suitable to the ALERT environment, but all failed because their architecture depends on a connection oriented bi-directional link. These include the ETS-300.230 air interface protocol, designed for reliable data transmission over private mobile radio (trunked systems) using convolutional coding FEC (rate ¾).[4] It was developed by the User Access Definition Group (UADG) for interoperability among different vendors of MPT1327 trunked radios. A competing protocol is Motorola's iDEN for trunked radio (including consumer telephone/wireless service). The third is Terrestrial Trunked Radio (TETRA) protocol, an open digital trunked radio standard defined by the European Telecommunications Standardization Institute (ETSI) for mobile radio users. Although TETRA provides connectionless service mode, its PSK modulation technology is incompatible with the ALERT requirement for a limiter discriminator based non-coherent demodulator (see the following Modulation Technology section).

Paging protocol requirements seem to be similar to ALERT requirements: The need for efficient, robust communications between low-cost, low-power devices sharing a single network. The primary paging protocols are Motorola's FLEX (and the two-way ReFLEX), [5] POCSAG and

---

[3] As shown above, many telemetry radio manufacturers design their own proprietary physical layer packet protocol.
[4] See Mobile Access Protocol for MPT1327 for trunking radios and MAP27protocol. (See http://www.teknologi.web.id/radiotrunked/mpt1327/map27.pdf , September 2002, for protocol description).
[5] Unfortunately, Motorola announced in 2001 that it will discontinue distributing its ReFLEX two-way paging protocol products to concentrate on the development of cellular network messaging products for GSM, GPRS and Code Division Multiple Access (CDMA). Some manufacturers still provide decoder integrated circuits, see HuneTec, Ltd., Korea, at http://www.hunetec.com/hunetec/product4.htm as of October 2002, but previous large suppliers, including Toshiba,

ERMES. POCSAG, developed in 1981, is limited to 1,200 bps. ERMES is the European 6,250 bps paging protocol.

Paging protocols are also unsuitable; their system architecture is designed for outbound one-way communications. The transmitter maintains complete control of the radio channel, and it is not battery powered or computationally constrained. To maximize the channel utilization it uses Time Division Multiple Access (TDMA) with continuous transmission. Individual packets destined for different pagers are concatenated and placed into specific time slots. These approaches are not applicable to ALERT, where there are multiple, independent transmitters sharing the same channel.

The most likely candidates for adaptation to our needs are from the two most widely adopted packet switched network protocols: DataTAC (ARDIS) and Mobitex. Their system architectures are the most similar to ALERT, and their physical layer packet structures are the most suitable.

ARDIS

This two-way radio service began as a joint venture between IBM and Motorola in 1983 to support dispatching, including public safety mobile data terminals. ARDIS was designed to exchange data files smaller than 10 Kbytes. It operates half duplex in the 800 MHz band, and initially provided 4800 bps in a 25 kHz channel using digital FSK and packet protocol named MDC-4800. The MDC-4800 protocol has been recently replaced by RD-LAP, which provides 19.2 kbps in a 25 kHz channel, using 4-Level GFSK modulation. The RD-LAP packet structure is shown in Figure 1.

Much of the desired ALERT packet structure is incorporated in the RD-LAP packet. Its minimum packet size is 21 data bytes plus 7 CRC error detection bytes, interleaved and convolutionally coded into a 56-byte packet. The packet size is flexible, expanding in 12-byte increments. The RD-LAP packet fails technically in two ways. Its preamble is only 48 bits long for bit sync and 50 bits long for frame sync, which is marginally too short to provide robust synchronization in very low S/N ALERT environments. Most significantly, ARDIS does not include Reed-Solomon forward error correction, without which any ASSP would not meet our bit error rate requirement.

National, and Philips are no longer providing ASSP integrated circuit solutions. See the press release at http://www.motorola.com/mediacenter/news/detail/0,1958,832_581_23,00.html as of October 2002.

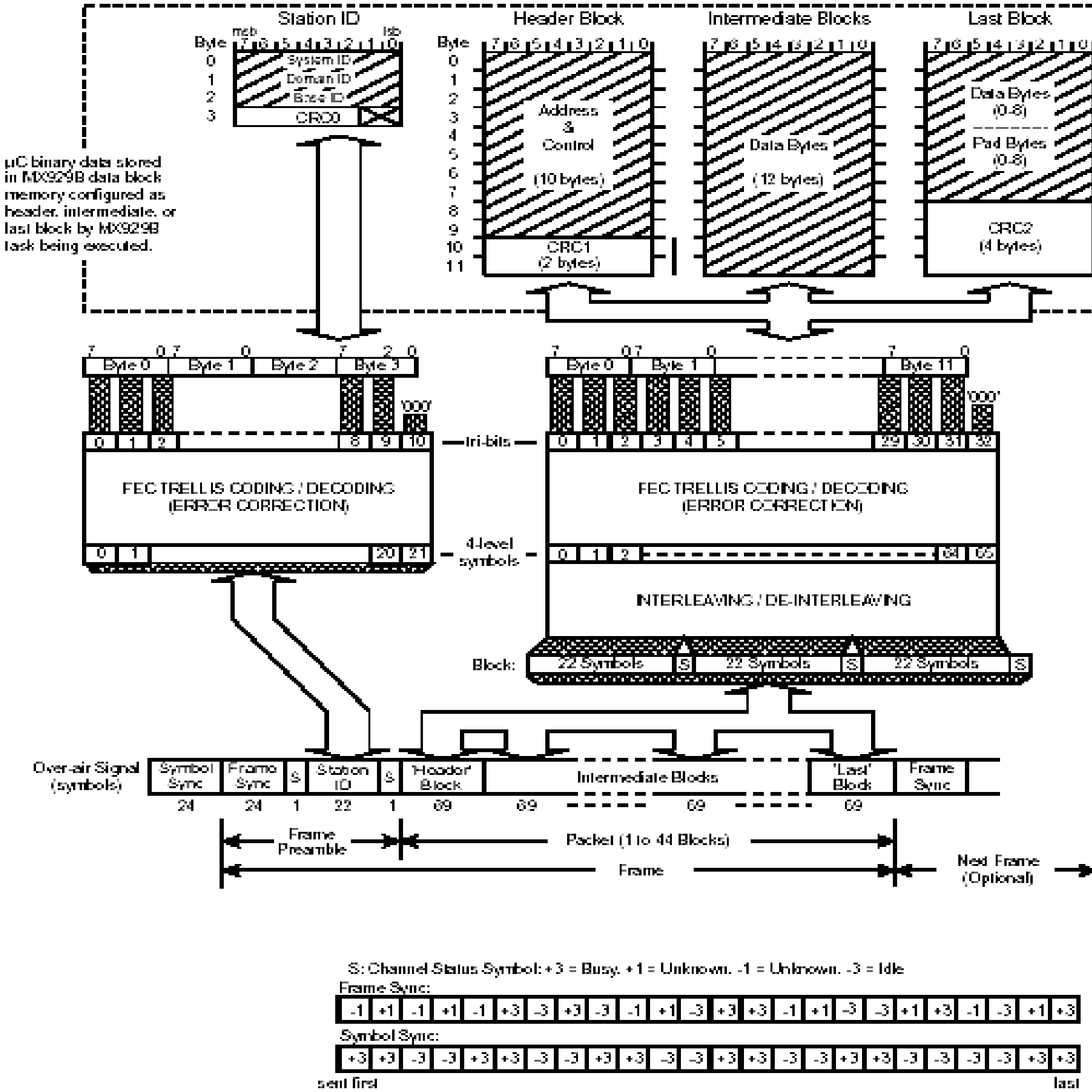


Figure 1: ARDIS Packet Structure (From CML-Microcircuits, Inc. MX929B specification sheet)

Mobitex

This is an interconnected trunk data network developed by Ericsson and Swedish Telecom that has been in operation in Sweden since 1986 and is deployed in more than 13 countries. It was designed for both voice and data services, but only data services are deployed in the US and Canada. It operates half duplex in 896-901 MHz and 935-940 MHz bands, providing 8000 bps in 12.5 kHz channels using GMSK modulation. The Mobitex physical packet structure is shown in Figure 2.

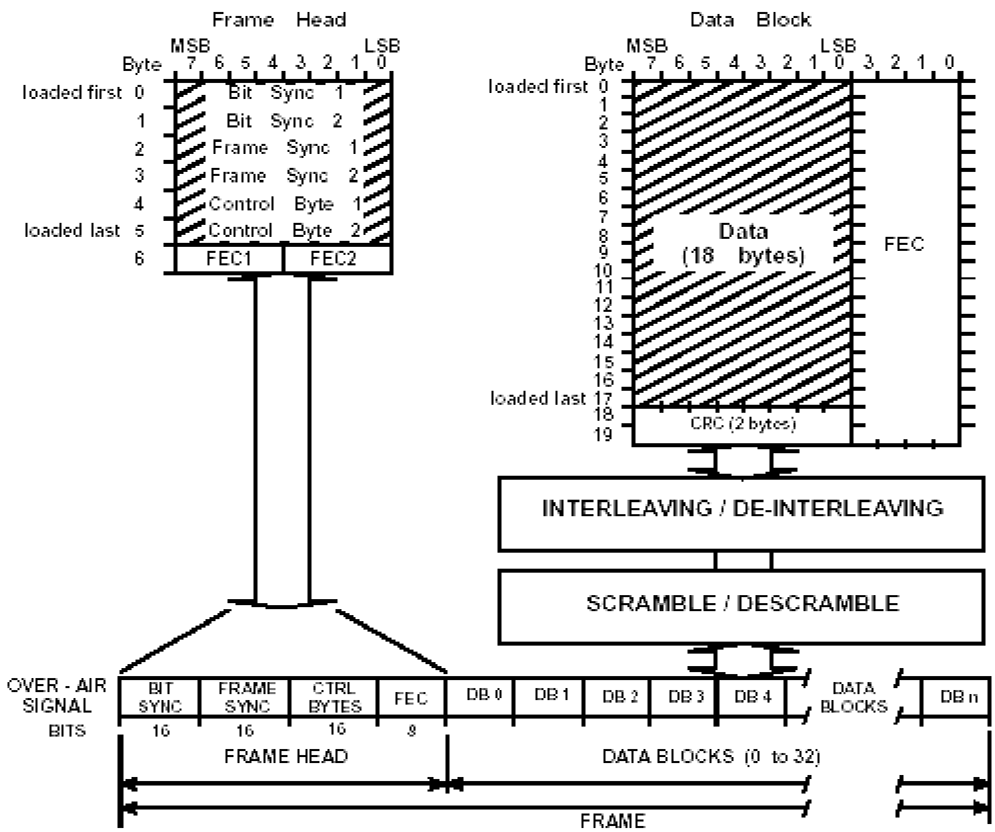


Figure 2: Mobitex packet structure (From CML-Microcircuits, Inc. MX909A specification sheet)

The Mobitex packet structure also incorporates some of the desired ALERT features. It has a minimum packet size of 37 bytes that contains 18 data bytes, 2 CRC bytes, 10 bytes FEC and 7 bytes of preamble. It uses FEC with interleaving and has a flexible total data packet size. Unfortunately, its minimal bit and frame synchronization (only 16 bits for each) and lack of convolutional coding limit its robustness and impair it as a direct fit for a new ALERT packet structure.

Another significant constraint for both packet switched protocols is their proprietary status. The ARDIS protocols (MDC-4800 and RD-LAP) are proprietary to Motorola and must be licensed. Similarly, the Mobitex protocol must be licensed from Ericsson Mobile Data Design AB. The licensing issue may not be insurmountable, depending upon price, but it would require all ALERT vendors to license the technology if they were to build their own modem. This complicating factor would discourage multi-vendor implementation.

Amateur Packet Radio AX.25

Another candidate protocol evaluated is the Amateur Packet Radio AX.25 (Amateur X.25) Link Access Protocol.[6] It defines both a physical layer and the data link layer of the ISO model; it is based on the 1970s X.25 synchronous link protocol. Implemented on half or full duplex radio circuits, the protocol is designed to establish a bi-directional connection between two stations, allowing multiple simultaneous connections.

[6] The full specification for Version 2.2 (1997), is available at Tucson Amateur Packet Radio Corporation at http://www.tapr.org/ as of October 2002.

Recently, APRS (Automatic Packet Reporting System), using a subset of the full AX.25 protocol has become popular. It utilizes the non-connection oriented "Unnumbered Information Frames" capability in the AX.25 specification to enable broadcast messages from units, similar to the broadcast messages of the current ALERT system. APRS was originally used to broadcast GPS position reports from a remote in order to provide continuous location updates to one or more receivers, but is now used to broadcast a variety of short packet multicast information.

The original 1970's X.25 protocol was first developed for wire-line synchronous linkage, and unfortunately AX.25 retains cumbersome legacy mechanisms like "bit stuffing". It also does not use any form of FEC channel coding. These limitations make it sub-optimal for the new ALERT protocol. Its frame sync, called a flag, exhibits an extremely poor auto-correlation response, making it virtually unusable for frame synchronization in low S/N environments. The data fields are ASCII encoded for simplicity of decoding, but this would be very inefficient for ALERT data. Its primary advantages are the availability of low cost AX.25 and APRS modems, complete protocol engines called terminal network controllers (TNCs) and significant open source software.

Results from Review of Existing Commercial Protocols

We were unable to find an existing public physical packet structure that combines all the features we seek. We believe we have reviewed the most representative and likely candidates for wholesale application to ALERT.

There are many more protocols for radio communications not discussed here, for instance: NEMA 1083, which is used in Marine, HF and meteor burst communications; MANET, an IP based intelligent sensor network; and APCO 25, a public safety standard for digital two-way radios. Given enough time and effort, one might be discovered that is adoptable, but we do not believe a more exhaustive effort would be cost-effective.

Our examination of existing commercial or standard radio network protocols shows that each has been optimized for the unique features of its environment, message type, and network configuration. The needs of the ALERT community represent another unique combination of these factors. It appears that a new ALERT protocol will require its own solution that incorporates many common principles and components of existing commercial protocols.

ALERT Channel Capacity and Channel Impairments

Channel impairments in the Land Mobile VHF and UHF RF bands are:

1. **Noise**, typically modeled as a random Gaussian power distribution – referred to as Additive White Gaussian Noise (AWGN). AWGN is inescapable in any communication channel and causes random bit errors at a rate dependent on the ratio of signal power to noise power in the receiver bandwidth (S/N).

2. **Burst noise**, modeled as noise that swamps the signal continuously for a short period of time, typically caused by transient RF signal sources, such as strong near-by radar pulses.

3. **Multipath effects**, including distortion and fading caused by the destructive interference from time lagged or phase distorted reflected or refracted signals. Multipath is highly dependent on geographic terrain, RF frequency and the movement of the transmitter, receiver and reflecting and refracting surfaces. Multipath contains both fixed and time

varying components depending upon the change over time of the reflecting and refracting surfaces.

AWGN and multipath are the two channel impairments that limit the increase in the ALERT protocol bit rate. Burst noise is bit rate independent. The following sections explain what impact AWGN and Multipath has on higher bit rate ALERT transmissions and what technologies may be used to mitigate their effects. In summary, AWGN is managed using improved modulation techniques and channel coding, while multipath is managed by limiting the maximum symbol rate.

Signal-to-Noise Ratio and Energy-per-Bit

One of the limiting constraints for both the current and any increased bit rate ALERT air interface is providing enough RF power to overcome AWGN. Shannon's 1948 landmark paper [Shan48] proved that a channel capacity is limited by channel bandwidth and power. The error free capacity in bps on a channel perturbed only with AWGN is given by C,

$$C = B^*\log_2 (1 + S/N)$$

where B is channel bandwidth in Hz, S is the average signal power in watts, and N is the average additive noise power in watts. For most RF modulation technologies, "channel bandwidth" equals ½ the RF bandwidth. Given the current ALERT RF channel bandwidth of 12.5 kHz and a nominal S/N of 12 dB (which is 15.8 in dimensionless S/N of Watts/Watts), the theoretical error free ALERT air interface channel capacity is:

$$C = ½ * 12{,}500^* \log_2 (1 + 15.8) = 25{,}500 \text{ bits/second}$$

What is notable is the relationship between bandwidth, S/N and channel capacity. Even at 0 dB S/N, where noise power is equal to signal power, the theoretical ALERT channel capacity is still 6,250 bps, error free! Although the theoretical limit is not achievable, practical systems have achieved within 0.7 dB of Shannon's limit with turbo-codes.[7] [Berro96]

Our goal is to increase the bit rate of ALERT without increasing the error rate. This requires understanding what effect increasing the bit rate has on the modulated signal. To do this, and to evaluate different modulation and channel coding methods, we need a figure of merit similar to the signal-to-noise ratio but normalized to eliminate the dependency on channel bandwidth and bit rate. The widely used figure of merit is Eb/N0: Eb is the energy per bit (in joules or watt-seconds) which is the average signal power, S (in watts) times the bit time, T_b in seconds; N0 is the noise spectral density (in watts/Hz) which is the average noise power, N divided by the receiver noise bandwidth, B_{RF}. The bit rate is R, the inverse of the bit time T_b. Then the Eb/N0 figure of merit is: [Sklr01][8]

$$Eb/N0 = (S^*T_b) / (N/B_{RF}) = (S/R) / (N/B_{RF}) = (S/N) * (B_{RF}/R)$$

The key concept is that when doubling the bit rate without changing any other parameters, Eb/N0 is halved, or decreased by about 3 dB. Intuitively, transmitting twice the number of symbols in a

[7] Turbo-codes were first described in [Berr96]. They are generated with multiple recursive convolutional coders operating in parallel on a data stream with their outputs linked together by non-uniform interleaving before being transmitted. Decoding is done iteratively, making use of information available from the other decoder from the previous step. The original paper demonstrated coding gains of 10.5 dB at BER of 10^{-6} for QPSK systems. See also Comtech Advanced Hardware Architecture Inc. at www.aha.com for Turbo Product Code application information. A general description is available in [Guma98]. Xilink Inc. sells a Turbo Code coder/decoder "logic core" for their Field Programmable Logic Array (FPGA) for cellular 3G applications (WCDMA) that achieves a BER of 1×10^{-6} at an Eb/N0 of 1.5 dB. See: http://www.xilinx.com/ipcenter/turbo_convol_lounge/

[8] See [Sklar01] page 117.

fixed time period with the same total RF energy halves the energy in each bit, while the noise power N0 remains constant, assuming the receiver noise bandwidth doesn't change.

Therefore, if the bit rate of the ALERT air interface increases from 300 bps to 2400 bps with no change in transmitter EIRP, receiver antenna or sensitivity, Eb/N0 decreases by 9 dB; changing from 300 bps to 3600 bps results in a decrease of 10.8 db; and from 300 bps to 4800 bps results in a decrease of 12 dB.

With no other changes, decreasing the energy in each bit will increase the statistical probability that noise power during a bit period will cause a bit error. For any digital modulation technology the theoretical ideal[9] probability of a bit error (Pe) can be expressed as a function of Eb/N0 alone.[10] For example, if the current ALERT system used a simple binary, non-orthogonal[11] frequency shift keying (FSK) modulation that was then non-coherently detected[12], the Pe[13] function is:

$$Pe = \tfrac{1}{2} * e^{(-\frac{1}{2} * (Eb/N0))}$$

The equation can be evaluated to show that to theoretically achieve a Pe of 1×10^{-4} (or an average of 1 bit error per 10,000 bits) requires an Eb/N0 of 12.3 dB; a Pe of 1×10^{-5} requires Eb/N0 of 13.4 dB.

Based on these facts, if we are to increase the ALERT bit rate without impacting the probability of error, then we must find a modulation technology that, in combination with channel coding methods, achieves an equivalent Pe at a lower Eb/N0. In order to increase to 3600 bps we must tolerate an Eb/N0 that is 10.8 dB lower, or at 2400 bps 9 dB lower, than the current Audio FSK FM modulation.

Theoretically it is possible to increase the ALERT air interface bit rates well beyond 3600 bps, with no bit errors in a re-farmed 12.5 kHz channel. While the full theoretical capabilities may not be achievable, advances in technology in the last 25 years make substantial increases to the ALERT bit rate practical and cost effective. The keys to this implementation are more power-efficient modulation methods and channel coding that permits recovery from errors introduced by noise.

---

[9] Except when explicitly noted, the relationship between Eb/N0 and BER always refers to the theoretical ideal. In practice the Eb/N0 required to achieve a specific BER is higher due to circuit noise and imperfections. The reasonable working assumption is that comparing the theoretical ideal Eb/N0 versus BER between differing modulation technologies translates to equivalent differences when realized in circuits, provided the demodulation circuit complexity is similar.

[10] The probability of bit error does vary for digital FSK depending upon demodulation technology: coherent, non-coherent or limiter discriminator. In all PSK and PSK hybrid modulation schemes, except the special case of pi/4 QPSK, the demodulator is required to be coherent. Digital AM is also possible to demodulate non-coherently, but is not a robust technology for ALERT use due to significant amplitude variations caused by the typical radio channel.

[11] Orthogonality for non-coherent detected FSK is when all frequency shifts satisfy $f1 - f2 = n/T_b$, where n is an integer and T_b is the symbol time; for coherent detected FSK it is when all frequency shifts satisfy $f1\text{-}f2 = n/(2^*T_b)$. See [Sklar01] page 110 for definition; page 129 for mathematical definition. See page 200-204 for definition of orthogonal FSK signal tones.

[12] Non-coherent means without the use of precise phase information; e.g. typical non-coherent FSK demodulation is done by using bandpass filters centered each frequency (determined a priori), followed by an envelope detector for each frequency. Coherent FSK demodulation is typically done using a reference oscillator, phase locked to the average received frequency, generating an in-phase (I) frequency signal and a quadrature (90 degrees out of phase, Q) signal, each then independently mixed with the received signal, the outputs of which are then sampled for phase shifts representing the frequency shifts. Alternatively, coherent detection is done by means of a "correlation receiver" which determines the best correlation, or match, between the incoming signal and a set of reference frequency signals, each generated from the received signal frequency. See HP Application Note 1298, page 15.

[13] See [Sklar01] page 219.

Modulation Technology: Bandwidth Efficiency versus Power Efficiency

A requirement of a new air interface is that it can be implemented by interfacing the modem to the audio input and output of off-the-shelf FM transceivers like those currently in ALERT use. This places several constraints on modulation technologies that can be considered.

Briefly, a FM transceiver[14] transmitter circuitry converts every unique voltage level at the audio input into a unique (RF) frequency at the RF output, typically with the 0 volts level creating a RF frequency exactly at the assigned radio channel frequency (e.g. 169.25 MHz).[15] As the voltage varies, the RF frequency varies proportionally, related by a modulation index (e.g. 1 V moves the frequency by 3 kHz, to 169.253 MHz). This amplitude modulation to frequency modulation conversion is typically implemented with a Voltage Controlled Oscillator (VCO). Conversely, the receiver circuitry amplifies the signal detected by the antenna, filters it to the band of interest (e.g. a 25 kHz bandwidth centered on 169.25 MHz), then with what is called a limiter discriminator, converts the instantaneous RF frequency into a voltage at the receiver output (e.g. with 169.253 MHz received, 1 V is output). A change at the audio input cannot change the amplitude of the RF carrier, nor can it directly modify the phase of the RF carrier. Similarly, the receiver's output does not vary depending upon variations in signal amplitude, nor does it respond directly to instantaneous changes in the phase of the RF carrier.

Obviously, amplitude modulation schemes such as quadrature amplitude modulation (QAM) are impossible to implement with FM transceivers. Phase shift keying (PSK) modulation schemes, whether binary, quadrature or higher complexity, require modulation designs that can precisely control the phase of the RF carrier (e.g. shift the 169.25 MHz carrier 180 degrees to represent a binary value) and typically require demodulation designs that directly detect phase shifts (decode a binary value on a 180 degree phase shift). These are also impossible to implement using audio inputs and outputs of FM transceivers. Even Pi/4DQPSK modulation, which can be non-coherently, limiter discriminator demodulated, is not a good choice for the new ALERT because it is impossible to implement using solely an audio input to an FM transceiver. For the same reasons, none of the hybrid AM - PSK modulation technologies can be used for a new ALERT protocol.

In fact, it quickly becomes clear that because of the limitation of using an FM transceiver for ALERT, the only feasible modulation alternative is a form of frequency modulation. Figure 3 is a visual representation of the fundamental digital modulation technologies. One benefit of using FM is that since all the information is carried in the frequency, using a constant amplitude RF envelope, efficient 'hard-limiting" amplifiers can be used and radio channel induced variations in amplitude do not affect the information.

---

[14] Of the type currently and planned to be used for ALERT.
[15] Assuming the audio input is DC coupled.

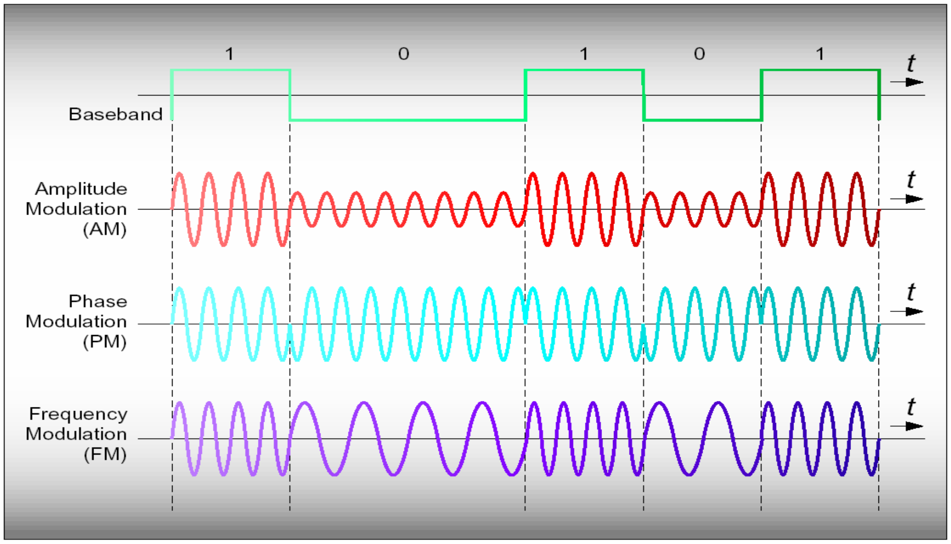


Figure 3: Simple Digital Modulation Schemes

A modulation technology's fundamental performance is typically measured by two characteristics. The first, bandwidth efficiency, relates to how much information (bps) can be transmitted per Hz of bandwidth. The second, power efficiency, defines the signal-to-noise ratio required to maintain a certain probability of error, and is typically normalized to Eb/N0. When various modulation technologies are plotted on a chart of bandwidth efficiency versus power efficiency, their relative attributes can be easily compared.

Figure 4 is a plot[16] of bandwidth efficiency versus power efficiency for common digital modulation technologies. The horizontal axis is power efficiency – the required Eb/N0 for a probability of bit error equal to 1×10^{-5}. The vertical axis is bandwidth efficiency – the bits/sec/Hz capability of the modulation. The sloping line from the top right to the lower left is Shannon's limit. The ideal modulation technology is the one furthest to the left (best power efficiency) and highest (best bandwidth efficiency).

The family of FSK modulation schemes, whether demodulated coherently, non-coherently, or limiter discriminator appears at the lowest portion of the graph: they exhibit the lowest bandwidth efficiencies. Binary FSK modulation (M=2)[17] has the worst power efficiency of any other modulation technology, except for some M=16 technologies, when measured in terms of their distance from Shannon's limit. Yet, digital FSK frequently is the best option due to 1) its constant RF envelope, 2) its demodulator flexibility – it can be demodulated using a complex coherent, a less complex non-coherent, or a simple, low cost, sub-optimal limiter discriminator demodulator, and 3) its ability to utilize off the shelf transceivers.

Figure 4 clearly illustrates why PSK is usually used when the channel is band limited. Using M=8 Continuous PSK (CPSK), more than 6 times the information can be transmitted in the same channel at the same Eb/N0 power efficiency as non-coherently demodulated binary FSK.

[16] From [Bened87] page 482.
[17] M is the number of states for each symbol or baud. For binary modulation schemes M=2 and each baud conveys a single bit of information (and hence baud = bps). For M > 2 baud rate does not equal bits per second: baud rate = $\log_2(M)$ bits per second; e.g. M=4 (such as QPSK), each symbol transmitted (baud) conveys 2 bits of information. M-ary is a generalized term for multiple level modulation, i.e. where M > 2. Quaternary refers to M=4.

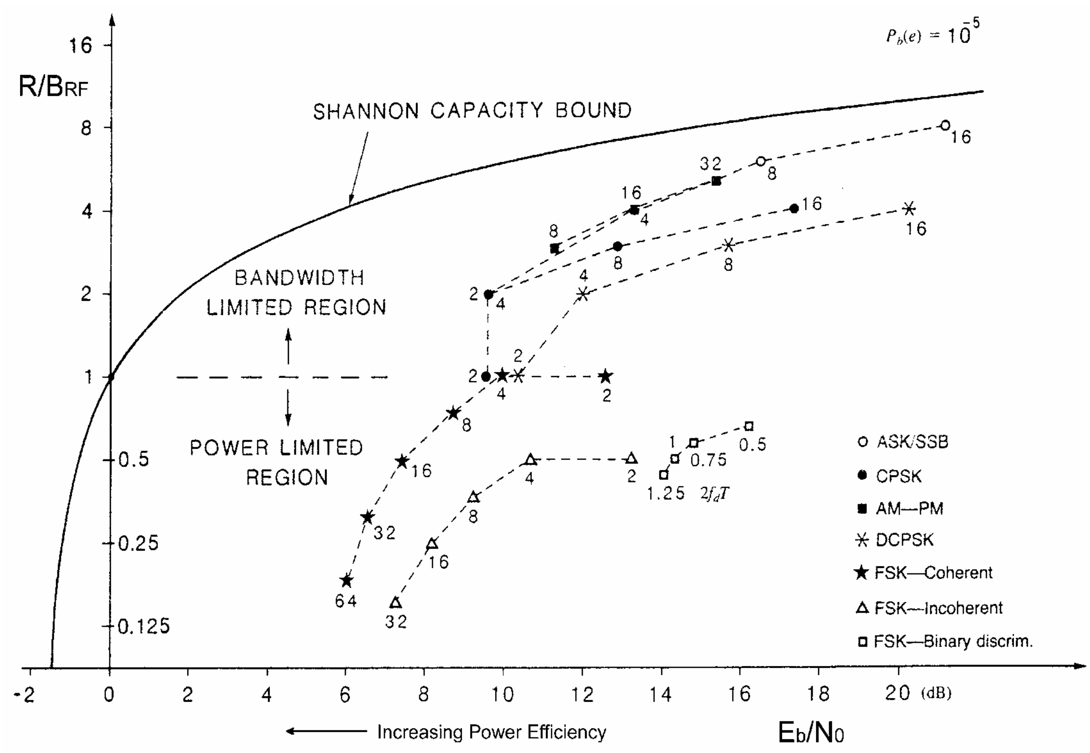


Figure 4: Digital Modulation - Bandwidth Efficiency versus Power Efficiency

To propose an improved modulation technology requires understanding the performance of the current ALERT modulation technology. Unfortunately, the current ALERT modulation is not represented by any point on Figure 4, since it is not a digital modulation technology. Digital modulation modifies one of the fundamental components (amplitude, frequency or phase) of the RF carrier in discrete increments to represent discrete values.

The ALERT protocol is technically Audio FSK FM (AFSK-FM): the digital information is carrier by using frequency shift keying at audio frequencies, which is then FM modulated. Two audio tones (time varying analog voltages, sine waves) are used to represent two binary values. This analog signal is then applied to an analog FM radio modulator. This creates an RF carrier whose frequency varies smoothly in time, at each instant exhibiting a different instantaneous frequency, directly proportional to the sine wave analog voltage at the input to the FM modulator. The RF carrier itself does not carry the binary information, nor does it change from one discrete value to another so therefore it is not a digital modulation technology at the RF channel.

Audio FSK is a legacy from data transfer over telephone lines. The ALERT modulation is an adaptation of the 1963 Bell System 103 modem standard, revolutionary at the time, which uses audio FSK (not RF modulated) to provide simultaneous bi-directional digital communications at 300 bps on voice grade telephone wires where it uses a frequency bandwidth of 300 Hz to 3000 Hz. Soon thereafter, it was adapted to provide half-duplex communications on radio links by utilizing available analog FM voice transmitters, since they also were designed for voice communications with approximately the same audio channel bandwidth. The Bell 103 standard accommodated bi-directional communications by defining 4 non-harmonically related audio tones that fit within the 2700 Hz audio bandwidth, two for each communication direction. Since simultaneous bi-directional communication is not possible in a single radio channel, in the late

70's the standard for radio telemetry (AFSK-FM) became the Bell 202 standard, using two audio tones spaced wider within the audio bandwidth, enabling 1200 bps.

Note that Figure 4's comparison of digital modulation performance is based on using an ideal RF modulator at the transmitter and an ideal RF demodulator at the receiver. The comparison uses a statistical analysis of the impact of AWGN added to the signal between the modulator and demodulator. When implemented, of course, due to imperfections in real circuits, no modulator or demodulator achieves ideal performance. Unfortunately, AFSK-FM demodulation requires two conversions before a decision detector (whereas digital FSK requires only one). In AFSK-FM demodulation the RF FM is converted to an analog waveform, typically with a limiter discriminator, the output of which is used by a second demodulator, the audio FSK demodulator, to generate an analog voltage to be tested to decide the binary value. Hence, circuit realizations of AFSK-FM demodulators require two imperfect circuits prior to the decision circuit, degrading its performance from its ideal. Additionally, in AFSK-FM, the RF FM demodulator is used in a linear mode – to recreate the audio tone. Any distortion or non-linearity of the RF FM demodulator decreases the S/N of the audio signal, which is the input to the audio FSK demodulator.

An analytical model for AFSK-FM BER performance can only be defined if a theoretical model for the "channel" between the AFSK demodulators is defined. But this "channel" model must include the RF FM modulator and demodulator performance – a difficult task unless it's assumed to be perfect, due to the variations in design and implementation. The modeling problem is best described by Tjhung: "Due to the non-Gaussian nature of the post discriminator noise, it is difficult to find the exact amplitude distribution of the noise in the output of the post discriminator filter." [Tjhu70] The authors have been unable to find a theoretical analytical model for the AFSK-FM system bit error rate, even with a thorough search in the Institute of Electrical and Electronic Engineers (IEEE) periodicals back to 1950 (the earliest available online).[18]

Without any AFSK-FM analytical data, the best alternative is to use the available empirical data to compare power efficiencies in implemented AFSK-FM and digital FSK systems. The empirical data is plotted on Figure 5, showing measured RF input levels (in dBm) versus BER performance, gathered from the sources described below. No source for 300 bps AFSK-FM performance information was discovered.

In 1983 Goode tested the Tucson Amateur Packet Radio's (TAPR) Terminal Network Controller & 1200 AFSK FM Modem (TNC) to determine its BER. [Goode83] The bench test setup used an HP RF Signal Generator and a Motorola Syntor VHF FM receiver in a loopback configuration through a single TNC-1200 bps modem. The measured RF input (dBm) versus BER is shown on Figure 5, labeled as "Gde 1200 AFSK".

Data from bench tests for radios using Amateur Radio's de facto standard GRUH raised cosine filtered binary FSK 9600 bps modulation is shown of Figure 5. The Amateur Radio Relay League (ARRL) Lab designed a standardized bench test to measure amateur radio manufacturers' transmitter and receiver performance for G3RUH 9600 bps operations. [Bloom95a] Performance for five transceivers: three UHF radios designed for voice, a VHF/UHF multimode radio designed for voice, and a UHF data-only radio was reported by Bloom. [Bloom95b]. The measured RF input (dBm) versus BER is also plotted on Figure 5 for the best radio (labeled "Blm 9600 RCFSK 1") and the worst radio (labeled "Blm 9600 RCFSK2") tested.

---

[18] C. S. Weaver, in "A Comparison of Several Types of Modulation," (see *IRE Transactions on Communications Systems*, pp. 96-101, March 1962) uses the mean square error to compare the BERs of PCM-FM (an early name for digital FSK) to "analog FM," since "…an analog voltage [from an FM demodulator] in general has no [analytical] error probability." He then proceeds to compare the mean square errors between FSK and analog FM, but does not then apply the results to determine an AFSK data transmission BER.

Additional data points are included for the Microwave Data Systems' MDS 4300 series of UHF data transceivers.[19] The MDS 3400 is not only selectable to operate at either 1200 bps AFSK-FM, 4800 bps digital FSK or 9600 bps digital FSK, but also the RF input level required for a 10^{-6} BER in each mode is specified by the manufacturer. Because the RF circuitry is common between all modes, and using the reasonable assumption the manufacturer has designed the MDS 4300 to optimize performance in all modes, these data are representative of the current state-of-the-art. The 3 data points are labeled "MDS …" on Figure 5.

Similarly, RF DataTech's QRT[20] UHF transceiver plus modem is selectable for 1200bps AFSK-FM and 4800 bps Gaussian filtered digital FSK operation and RF DataTech has specified the respective RF input levels required for a 10-3 BER. These 2 points are labeled "DT …" on Figure 5.

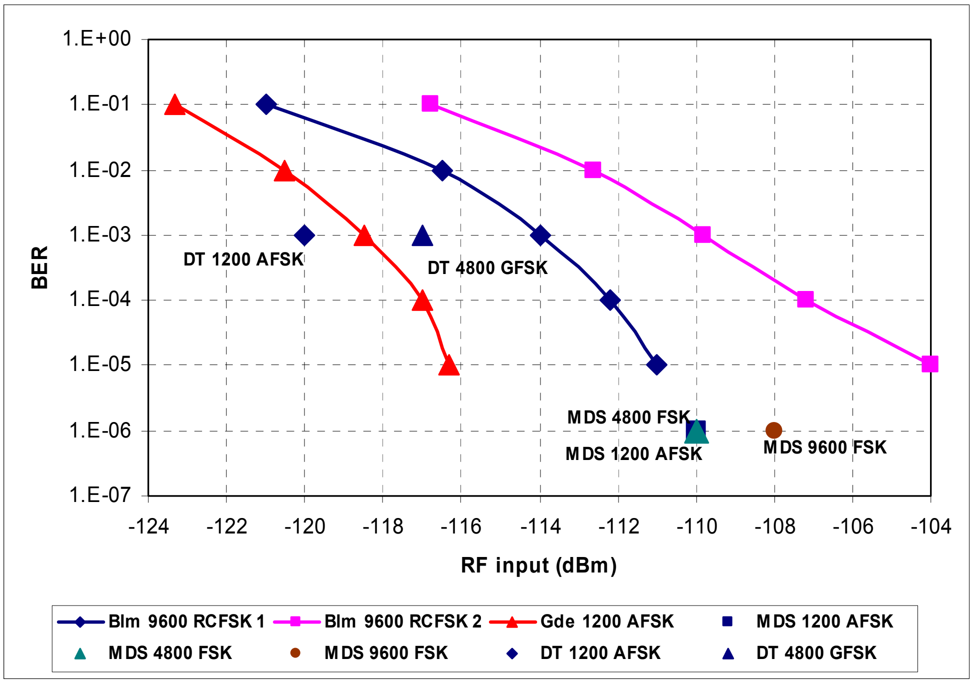


Figure 5: 1200bps versus 9600bps BER data

Figure 5 clearly shows significant performance variation for practical implementations of modems and RF transceivers. For instance, the MDS4310 is specified with identical BER at either 1200bps AFSK-FM or 4800 bps FSK, whereas the DataTech QRT shows a 3 dB difference between its 1200 bps AFSK-FM and 4800 bps GMSK. Also of interest, if the DataTech QRT 4800bps GFSK data point were extrapolated to the 10^{-6} BER level, using the curve shape of the Blm 9600 RCFSK1, it appears the DataTech QRT achieves 3 dB better performance than the MDS 4300 series. Obviously, the test setups were different, although it is reasonable to assume each attempted to replicate realistic operating environments. The radios used in the Bloom test represented early 1990's Amateur Radio commercial state-of-the-art, whereas the Goode test used an early 1980's commercial state-of-the-art.

[19] The MDS 3400 series specification sheet is available at: http://www.sine-wave.com.au/MDS4310.htm, as of December 2002.

[20] The Wireless Data Products, RF DataTech QRT modem specification sheet is available at: http://www.wireless-products.dk/PDF-filer/RF-DataTech-PDF/WP-QRT-Radio-modem.pdf, as of October 2002.

To normalize this empirical data and derive an Eb/N0 for 10^{-5} BER for AFSK-FM and digital FSK modulation, the 12 dB SINAD level and the receiver noise bandwidth for each radio is required. Additionally, since the 10^{-5} BER rates were not available for all radio-modems on Figure 5, where needed, the 10^{-5} BER points were estimated by extrapolation (modeled on the Blm 9600RC FSK1 curve). Table 1 shows the relevant parameters[21] and resultant calculated Eb/N0 for each of the systems.

Radio - bps - Modulation	12 dB SINAD (dBm)	Channel BW (kHz)	S/N for 10^{-5} BER	Eb/N0
Blm 9600 RCFSK 1	-113.0	25.0	13.8	18.0
Blm 9600 RCFSK 2	-108.8	25.0	16.7	20.9
MDS 9600 FSK	-119.0	25.0	22.0	26.1
MDS 4800 FSK	-119.0	12.5	20.0	24.1
DT 4800 GMSK	-119.0	12.5	16.9	21.1
Gde 1200 AFSK	-121.0	25.0	24.7	37.9
MDS 1200 AFSK	-119.0	12.5	20.0	30.1
DT 1200 AFSK	-119.0	12.5	13.0	23.2

Table 1: Empirically Derived Eb/N0 – digital FSK versus AFSK-FM

Again, there are significant variations within each modulation type, representing different RF and modem designs, and differing test setups. The filtered digital FSK modulation Eb/N0 spans an 8.1 dB range, while AFSK-FM spans a 14.7 dB range.[22] Some variation in the digital FSK result is to be expected due to the different forms of baseband filtering used in the equipment tested.[23] Inspection shows that the MDS radio consistently under-performs at all bit rates: from 3 to 6 dB worse than any other radio, except for the Syntor radio used in the Goode tests at 1200 bps. Note also that the best performing FSK modulation has an Eb/N0 of 18.0 dB, approximately 4 dB worse than theoretically predicted from Figure 4 for a limiter discriminator demodulation of FSK. Although there is insufficient data for accurate statistical analysis, the average and best performance for each modulation technology are shown in Table 2 below.

	AFSK-FM Eb/N0 (dB)	Filtered digital FSK Eb/N0 (dB)	FSK Improvement (dB)
Average	30.4	22.0	8.4
Best Performer	23.2	18.0	5.2

Table 2: Digital FSK versus AFSK-FM

Table 2 clearly illustrates that digital FSK exhibits a power efficiency improvement of approximately 6 dB compared to AFSK-FM.

We have derived an Eb/N0 for AFSK FM using empirical data for a 1200 bps data rate. Without an analytical model for the audio demodulator we can make the reasonable assumption that,

[21] A simplification that dB quieting (dBq) = S/N was used for the Gde 1200 AFSK-FM data, since SINAD was not provided. dBq is not precisely equal to S/N, since dBq is measured using an unmodulated carrier only. S/N is calculated from SINAD and used to determine Eb/N0.
[22] The receiver noise bandwidth was not ascertainable from the Goode 1200 bps AFSK-FM test setup description, therefore 25 kHz was assumed, since in the early 1980's 25 kHz channel usage was the most common; using 12.5 kHz would reduce the 37.9 dB Eb/N0 to 34.9, more similar to the data.
[23] Digital FSK baseband filtering is described in the next section.

given the same Eb/N0 at the audio demodulator input, the output bit error rate is the same regardless of the bit rate. Hence the Eb/N0 for 300 bps and 1200 bps AFSK-FM are the same. Using a form of digital FSK modulation for the new ALERT protocol will provide a 6 dB improvement in Eb/N0 at the same error rate. This allows the new ALERT protocol to transmit at 1200 bps with the same probability of error, in AWGN channels, as the current ALERT 300 bps.

Continuous Phase Modulation, Minimum Shift Keying and Bandwidth

Numerous variations of digital RF FSK have been mathematically described, simulated and practically implemented in the past 40 years.[24] Research has focused on both improving the power efficiency and bandwidth efficiency, but due to the high cost of spectrum, more research has targeted maximizing bandwidth efficiency, which is only a secondary concern for the new ALERT protocol.[25]

Minimizing the rate of change of the RF carrier's phase transitions improves FSK's bandwidth efficiency. The power in the "sidelobes"[26] of a FSK's power spectral density (PSD) is directly proportional to the rate of change and higher order derivatives of the FSK's phase transition function. The signal driving the frequency shifts, and therefore the phase shifts, is called the "baseband." For unfiltered FSK, the baseband is the square wave representing the binary data. Filtering the baseband slows the transitions, which then smoothes the rate of change of the phase transitions of the carrier. By lowering the sidelobes, bandwidth efficiency is significantly improved. Technically, since the PSD is the Fourier transform of the baseband signal, the unfiltered binary square wave baseband of FSK creates a PSD that is the sum of two sinc waveforms, with infinite sidelobes. FSK without phase discontinuities is termed continuous phase FSK (CPFSK), or more generally continuous phase modulation (CPM).[27]

Another method to maximize bandwidth efficiency is to minimize the total RF carrier frequency shift. For digital FSK, the relationship between the RF carrier frequency shift and bit period (T_b) is termed the modulation index, h. Expressed mathematically, the modulation index is $h = (f_1 - f_2) * T_b$. The minimum frequency shift that also forces transitions to occur only on zero crossing, minimizing the rate of phase transition, depends directly on bit rate. For non-coherent demodulation, orthogonal signals are frequency shifts separated by the bit rate, i.e. $\Delta f = n/T_b$, where n is an integer. So the minimum non-coherently demodulated frequency shift is $\Delta f = 1/T_b$.

With coherent demodulation, where accurate phase information is utilized to decode the signal, the minimum frequency shift that provides phase continuity and is "coherently orthogonal," is only one half the bit rate, i.e. $\Delta f = (1/2) * (1/T_b)$, hence it is FSK in which $h = {}^1/_2$. Since this frequency shift is the minimum orthogonal shift possible, it was termed Minimum Shift Keying (MSK), and was one of the earliest minimum bandwidth CPFSK modulation schemes studied. In fact, MSK is mathematically equivalent to sinusoidal baseband filtered Offset QPSK (OQPSK). Therefore, using a coherent receiver,[28] MSK achieves the same power and bandwidth efficiency as QPSK[29] with the added benefit that the bit synchronization signals are readily recovered by direct manipulation of the received signal. As an example, 2400 bps binary MSK requires frequency

---

[24] See [Sund86] for a comprehensive overview as of 1986.
[25] With the very short data content of simple ALERT messages, increasing the bit rate has diminishing returns for increasing channel capacity, due to fixed bit lock and frame synchronization time. A new ALERT protocol operating at 4800 bps in a 12.5 kHz channel would require a bandwidth efficiency of only 0.4 bits/sec/Hz, well within unfiltered FSK capabilities.
[26] "Sidelobes" refers to all the peaks after the first null in a PSD.
[27] To be precise, CPFSK (continuous phase frequency shift keying) is a subclass of CPM where the instantaneous frequency is constant over the symbol period, although use of CPFSK has faded in the last decade in most journals; currently CPM is used to describe most continuous phase FSK systems.
[28] For coherent demodulation of any FSK signal, and specifically for MSK, the modulation must also be coherent, that is, where the RF carrier is directly manipulated, typically by I and Q modulation.
[29] Although not identified on Figure 4, the MSK is identical to the "CPSK M=4" point.

shifts of only 1200 Hz, i.e. the RF carrier is shifted +/-600 Hz about its center frequency. MSK, as with any FSK signal, can be non-coherently demodulated, although it sacrifices the 3 dB power efficiency of being a type of OQPSK. It also can be demodulated with a limiter discriminator, but since a limiter discriminator is a sub-optimal non-coherent demodulator, MSK suffers the additional degradation in power efficiency that any FSK suffers with a limiter discriminator detector.

Although MSK doesn't significantly improve power efficiency over other forms of FSK when limiter discriminator demodulated, MSK improves the bandwidth efficiency of FSK. The 99% RF bandwidth[30] of MSK is 1.2 times the symbol rate, making it significantly more bandwidth efficient than FSK. [Pasup79] For instance, the 2400 bps MSK example has a bandwidth of only 2.88 kHz. MSK's minimum bandwidth characteristic is the reason it is used in many commercial applications.

Baseband filtering primarily is used to improve bandwidth efficiency. Two major classes of baseband filtering exist – "full response" CPM, characterized by baseband pulse shapes which are zero outside the symbol period, and "partial response" CPM, characterized by baseband pulses that spread over multiple symbol periods. The theoretical power efficiency of CPM is determined by a number of factors. For "full-response" CPM, the pulse shape, the modulation index and bits per symbol are the primary influences. [Aulin81a] For "partial response" CPM, an added factor is the baseband pulse symbol length. [Aulin81b]

CPM can be demodulated by monitoring the cumulative phase of the carrier. During modulation each symbol adds a unique amount of total phase shift total during its symbol period. These phase additions may occur linearly (from a rectangular baseband pulse shape), or non-linearly (from other than rectangular baseband pulse shape). At any point in time, the cumulative phase of the received carrier depends on the entire past history of the data transmission. By using partial response CPM with baseband pulses that extend over 3 to 4 symbol periods Eb/N0 power efficiencies can be increased by more than 4 dB compared to MSK. [Ander91] This requires coherent demodulation to take advantage of the information contained in the cumulative phase shift of the carrier. The more symbol periods over which each bit is spread, and the longer the observation time during demodulation, the better the performance. This gain, however, requires an exponential increase in demodulator complexity. [Iwan95] Sundberg describes mathematically the baseband pulse shaping required for L-Raised Cosine (RC), L-Spectral Raised Cosine (also called square root or SQRT RC), Tamed Frequency Modulation (TFM), L-Rectangular Pulse (REC) and Gaussian MSK (GMSK). In these designations, "L" refers to the length of the symbol pulse expressed in terms of symbol periods. [Sund86]

We do not propose to use partial response CPM for the new ALERT protocol since the necessary criteria can be met without the complexity of maximum likelihood sequence decoding. Moreover, without coherent demodulation the ALERT protocol cannot take full advantage of the increased power efficiencies of partial response CPM.

Filtering FSK – Cosine, Gaussian - GMSK

By using non-rectangular baseband pulses with FSK or MSK, significantly higher bandwidth efficiency is achieved with only a small degradation in power efficiency. The two most frequently adopted[31] pulse shapes are Gaussian and Raised Cosine. These baseband pulse shapes are

[30] There are many definitions of bandwidth; the important definition for FCC licensing is the bandwidth of the power spectral density containing 99% of the signal energy.
[31] In addition to the systems described in Appendix A that have adopted GMSK and GFSK, in 1995 the IEEE adopted Gaussian filtered FSK as the 802.11 wireless LAN modulation standard for 1 Mbps Frequency Hopping Spread Spectrum. In July 1997, IEEE 802.11 was adopted as a worldwide International Standards Organization (ISO) standard. The standard consists of three possible physical (PHY) layer implementations and a single common medium access control

implemented by low pass filtering the rectangular data signal. FSK with Gaussian baseband pulse shapes is GFSK, and MSK with Gaussian baseband pulses is GMSK.

The shape of the Gaussian and Raised Cosine filter are similar, and each is characterized in terms of the symbol period, although slightly differently. For Gaussian, the amount of filtering is defined by the dimensionless quantity BT, which is the -3 dB bandwidth frequency symbol period, or BT = $B_F * T_S$, where B_F is the filter -3 dB roll-off frequency and T_S is the symbol period. Similarly, for Raised Cosine, the -3 dB frequency point is set at ½ the symbol rate and the amount of filtering is dependent on a shape factor, referred to as α. When α = 0, the filter shape is a gentle slope, with a very wide resulting PSD bandwidth; whereas when α = 1 (physically unrealizable) the filter slope is vertical (called "brick wall"). Figure 6 shows the impulse response of Raised Cosine filter in the time domain for various shape factors (from [Valen01]).

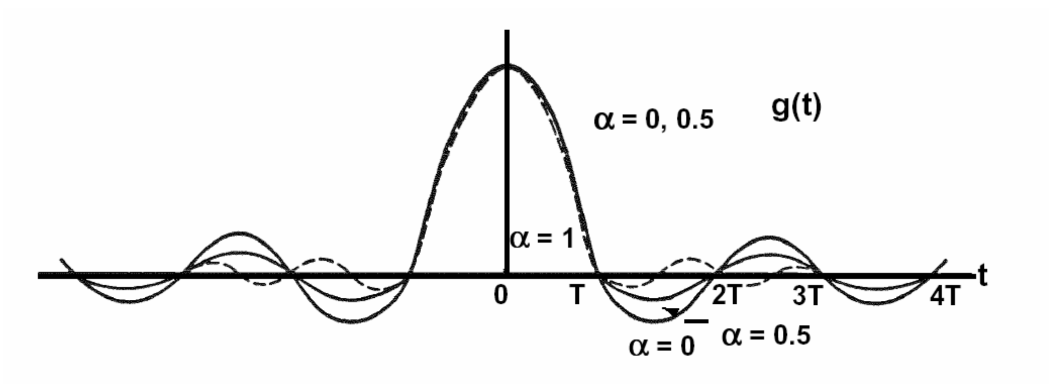


Figure 6: Raised Cosine Filter Impulse Response

For Gaussian filtering, decreasing BT increases bandwidth efficiency at the expense of increasing ISI, and increasing BER. Early studies demonstrated the degradation of BER for Gaussian filtered MSK was at a minimum at a BT of 0.587 at 1.4 dB.[32] [Ishiz80] Inspection of Figure 7 (taken from [Finton]) clearly demonstrates the effect of filtering, showing the PSD for the BT values of 0.3, 0.5 and infinity (GMSK with an infinite bandwidth Gaussian filter is MSK), compared to the power spectral density of QPSK.

(MAC) layer supporting data rates of 1 Mb/s or 2 Mb/s. The alternatives for PHY layer in the original standard include a frequency hopping spread spectrum (FHSS) system using 2 or 4 level Gaussian frequency-shift keying (GFSK) modulation and a direct sequence spread spectrum (DSSS) system using differential binary phase-shift keying (DBPSK) or differential quadrature phase-shift keying (DQPSK) baseband modulation.

[32] Using coherent detection versus BPSK.

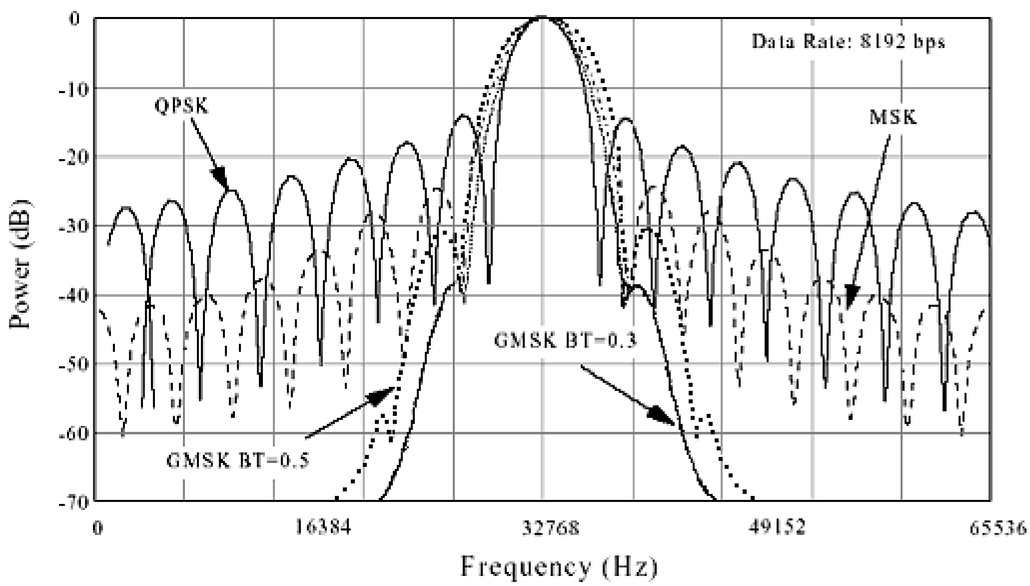


Figure 7: Comparison of Bandwidth Efficiency

Figure 8 characterizes the bandwidth occupancy of GMSK at h=1/2 for various values of BT. The x-axis shows the bandwidth, in Hz /bits/sec, and the y-axis shows the percentage of power contained within the specified bandwidth. The plot can be used to determine the required bandwidth for any particular bit rate at any of the plotted BTs. For example, the plot for BT = 0.5 shows that 99% of the power is contained within an RF bandwidth that is 1.04 times the bit rate. This indicates that a 4800 bps channel has a 99% bandwidth of slightly less than 5 kHz (from [Murat81]).

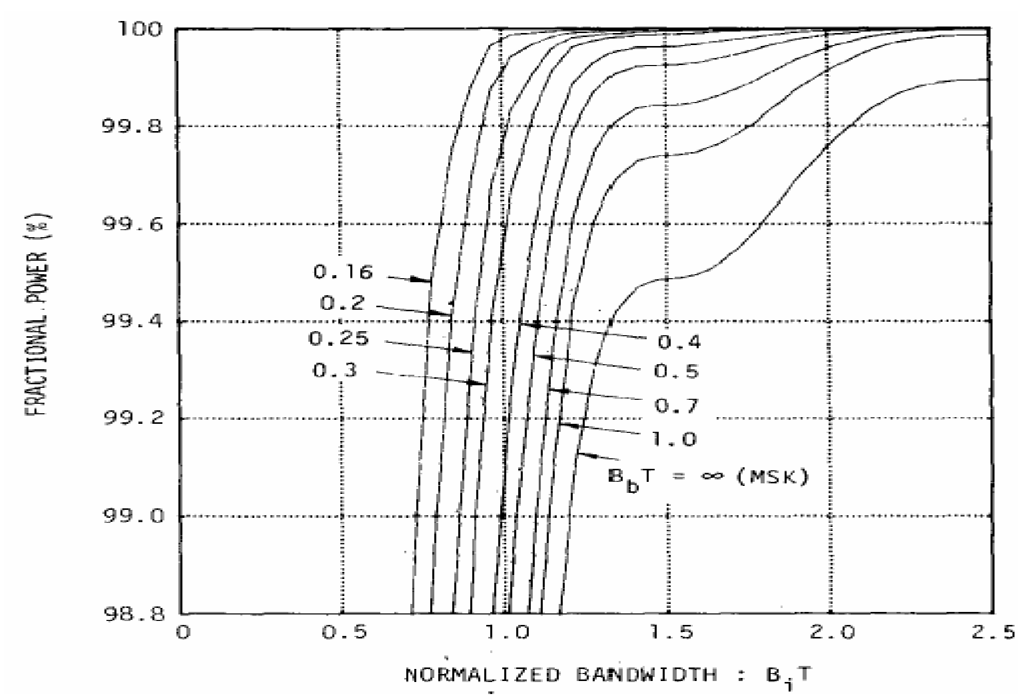


Figure 8: Bandwidth Occupancy of GMSK

Raised Cosine filtering provides similar bandwidth minimizing benefits to FSK or MSK. The trade off is that Gaussian filtering results in a non-Nyquist pulse shape, and therefore causes more inter-symbol interference (ISI).[33] On the other hand, Raised Cosine is Nyquist and therefore minimizes ISI, but is difficult to realize, requires truncation[34] and may cause ringing in practical hard-limited amplifiers. In the Amateur Radio community, terminal network controller modems using GFSK and RCFSK successfully interoperate with the pulse shape mismatch causing only a few dB degradation in performance.

Based on the research described above, the authors conclude that filtered digital FSK modulation is a suitable technology for ALERT. It provides approximately 6 dB of improvement in Eb/N0 over AFSK-FM, and a significant reduction in bandwidth. Additional research has shown power efficiency is maximized with a modulation index, h, of 0.7, and efficiency is also improved when BT is increased to 1.0-1.2. These changes increase occupied bandwidth, which, although not a concern for operations with 12.5 kHz channels at 4800 bps, would definitely not allow 4800 bps operation in a 6.25 kHz channel. Designing for a 6.25 kHz channel, even with BT=0.5, in practice limits the bit rate to approximately 4800 bps or less.

To enhance multipath resistance (discussed below) a good alternative may be four-level filtered FSK at 2400 symbols/second, which achieves a raw bit rate of 4800 bps. Using orthogonal frequency shifts, 4-FSK achieves better power efficiency at the same bandwidth efficiency. The increase in demodulator complexity does not, at first look, appear insurmountable; a number of available commercial VHF/UHF systems use 4-FSK to achieve increased bit rates.

Table 3 confirms that the proposed modulation technology is feasible. Created 10 years ago, it summarizes the significant research and analysis by the members of the IEEE 802.11 working group during a 5-year plus development effort to define the best “physical layer” for the wireless LAN protocol standard. [Chaya93]

---

[33] ISI is the effect of symbols adjacent in time interfering with each other. A Nyquist pulse shape is one that, although it may contain power in adjacent symbol periods, does not cause ISI because its power in adjacent symbol periods at the ideal sample time is zero.
[34] A Nyquist pulse is theoretically of infinite length in the time domain, so it must be truncated to be realized.

Modulation	GMSK	4-FSK	4-FSK	8-FSK	4-FSK
Originator	Motorola	N. Silberman	Proxim	Lannair	Lannair
Modulation Parameters	Gaussian filter BT=0.50 h=0.35	RC-2 h=0.25	Gaussian filter BT=0.5 h=0.1666	Pre-emphasis+ sqrt rolloff alpha=0.5 h=0.0833	Pre-emphasis+ sqrt rolloff alpha=0.5 h=0.1666
Reception method	limiter+ discriminator +postfiltering +2 lev. slicer	limiter+ discriminator +postfiltering +4 lev. slicer	limiter+ discriminator +postfiltering +4 lev. slicer	limiter+ discriminator +postfiltering +8 lev. slicer	limiter+ discriminator +postfiltering +4 lev. slicer
99% pwr BW	0.88 Fs	1.3 Fs	0.96 Fs	1.20 Fs	1.20 Fs
-20 dB BW	1.00 Fs	1.42 Fs	1.06 Fs	1.31 Fs	1.31 Fs
Symbol rate Conservative (Aggressive)	1,000 kb/sec (1130kb/sec)	700 ks/sec (750ks/sec)	960 ks/sec (1040ks/sec)	750 ks/sec (830ks/sec)	750 ks/sec (830ks/sec)
bits/symbol	1 bit/symbol	2 bits/symbol	2 bits/symbol	3 bits/symbol	2 bits/symbol
Bit rate Conservative (Aggressive)	1.0 Mb/sec (1.13 Mb/sec)	1.4 Mb/sec (1.6 Mb/sec)	1.92 Mb/sec (2.08 Mb/sec)	2.25 Mb/sec (2.5 Mb/sec)	1.5 Mb/sec (1.66 Mb/sec)
Adj ch interference	-26 dB	-26 dB	-25 dB	-25 dB	-25 dB
Alt ch interference	-78 dB	-70 dB	-71 dB	-74 dB	-75 dB
Es/N0 @ 1e-5	18.5 dB	20 dB	25.5 dB	26 dB	20 dB
Eb/N0 @ 1e-5	18.5 dB	17 dB	22.5 dB	21 dB	17 dB
C/N@1 MHz (coherent demod.)	18.5 dB	18.5 dB	24.5 dB	25 dB (22 dB)	19 dB (16 dB)

Table 3: Performance of Modulation Alternatives

In context of the new ALERT protocol, the 1993 chart demonstrates a number of points:

1. All the detection/demodulation methods were limiter discriminator.

2. Maximizing data throughput in a band-limited channel is emphasized by using small (or very small!) modulation indexes and BTs no greater than 0.5.

3. Even optimized for throughput, the power efficiency (Eb/N0 at BER of 10^{-5}) is only 4 dB down from MSK, and

4. The proposed modulation candidates for the new ALERT protocol are not “bleeding edge” – they are more conservative than any on the chart.

Available products

The feasibility of using Gaussian or Raised Cosine MSK or FSK (either binary or quaternary) for a new ALERT protocol is highly dependent on the development cost of a suitable modem. Unfortunately, the availability of a suitable ASSP implementing just the filtered MSK or FSK modulator/demodulator is limited by three factors:

1. Most modulator/demodulator functions today are included as portions of a highly integrated ASSP designed for a full protocol in applications such as Mobitex, Bluetooth or DECT.[35] A high level of on-chip integration minimizes overall system costs in high-volume products.

2. Much of the modulation/demodulation is done with software in digital signal processors (DSPs) or in the “logic cores” of field programmable gate arrays (FPGAs) with a proprietary hardware gate design. While the intellectual property (IP) of these cores may, in some cases be licensable, they rarely are open source.

3. Most filtered MSK and FSK modulation/demodulators are targeted to maximize bandwidth efficiency, at the expense of power efficiency.

For over 15 years, the Amateur Radio community has been operating high-speed filtered-FSK VHF and UHF data communications networks. Building on this existing public domain work can minimize the design costs associated with a new ALERT protocol modem. In 1988, Miller developed and released a 9600 bps modem design for Amateur Packet Radio, using binary encoded SQRT RC FSK, interfaced directly to the audio input/output of FM transceivers. [Miller88 and Miller91] Since that initial design, many other compatible implementations have been released based on FPGAs,[36] DSPs,[37] standard digital logic integrated circuits[38] and on Atmel’s low cost 8-bit AVR microcontroller (the G4XYW modem[39]). Either the DSP-based implementation or the AVR-based G4XYW could be adapted for use as a new ALERT 3600 bps to 4800 bps, RC or Gaussian filtered FSK or MSK modem.

A second option is to use to the CML Microcircuits CMX589A modem IC. It is a complete GMSK modem for wireless applications such as CDPD and Mobitex. It is low cost: $11.68 each[40] in 100 quantities, and low power: 1 mA at 3.0 V, and its performance specification is excellent: with a BER specification of $1x10^{-5}$ at S/N of 12 dB.[41] But it has three significant limitations if used for the new ALERT protocol:

1. Its filter constant is selectable only between BT = 0.3 or 0.5.

2. Its lowest specified bit rate is 4,000 bps; and

3. It is not capable of non-Gaussian filtering or quaternary encoding.

According to the manufacturer’s engineering department, the 4,000 bps specification is not actually a circuit limitation and the CMX will easily run at 2400 bps or 3600 bps.[42] It would be

---

[35] For example, the Philips PCF87750 ASSP is a single integrated circuit that implements a full Bluetooth 1.1 compliant protocol, including GFSK BT=0.5, h=0.3 modulation/demodulation, bit scrambling, FEC, CRC detection, power supply regulation and DC/DC conversion, with an ARM embedded controller, a multitude of peripherals and 384 Kbytes of flash memory, as well as voice codecs.

[36] See Yet Another Modem (YAM) at http://www.microlet.com/yam/ (as of May 2002), based on Xilinx's Xc5202, released in 1997 by Nico Palermo, IV3NWV.

[37] See Etxebarria’s “The new G3RUH software modem for the DSP56002EVM”, October 2000; or see Tucson Amateur Radio Packet Association’s (TARPA) DSP page at http://www.tapr.org/tapr/html/Fevm56k.index.html or http://www.dougbraun.com/evm.html as of October 2002.

[38] See John Magliacane, “The KD2BD 9600 Baud Modem,” available at: http://www.amsat.org/amsat/articles/kd2bd/9k6modem/index.html as of October 2002; originally published in the February, March and April 1998 issues of Satellite Times.

[39] The G4XYW modem information is available from the Thames Valley IP Users Group, at http://www.tvipug.org/download.html as of October 2002.

[40] Telephone conversation with CML Microcircuits sales on October 3, 2002; available in a plastic dual in- line package.

[41] The “best case transceiver interface”.

[42] Telephone conversation with Fred Banks in Engineering at CML Microcircuits October 4, 2002.

feasible to interoperate modems designed using the CMX 589 with other modem implementations as long as the new ALERT modulation was defined as GMSK with BT = 0.5.

Channel Coding – Convolutional and Block Error Correction Codes

Channel coding is the second technology we recommend using to sustain ALERT performance at a new higher bit rate. Channel coding, a form of forward error correction, nominally reduces data throughput by adding redundant bits into the data stream. However, the redundant bits improve the bit error rate of a channel to a degree that greatly outweighs the cost of their overhead.

There are two fundamental types of channel codes used: Trellis codes, the linear types of which are convolutional codes, and block codes. Convolutional codes reduce random bit errors using significant redundancy, while block codes reduce burst errors with a much lower overhead.

A third form of channel coding is interleaving, in which the data bits are reordered before transmission and returned to the proper order before decoding. No redundant bits are added, but the reordering further protects against burst events causing unrecoverable bit errors. Combining multiple codes, or multiple types of codes, enhances channel coding performance. When done sequentially, the resulting product is a concatenated code, and when done in parallel, it is known as a turbo code. Although turbo codes currently are used extensively[43] to provide performance near Shannon's limit, we do not recommend them for the new ALERT protocol for three reasons:

1. There is significant additional encoding and decoding complexity[44];

2. A turbo code normally is applied across a large data block, and is poorly suited to the small data content of a simple ALERT message;

3. Since turbo codes are new (1993), some codes and algorithms are patented. There is a potential for patent infringement or license issues. [CCSDS01]

One performance metric for channel coding is the coding gain[45], defined as the decrease in required Eb/N0 to achieve the same bit error rate in a channel perturbed only by AWGN. An alternative metric is the Hamming distance of a code, a mathematical property that defines the maximum number of bit errors a code can correct. Coding gain typically is used to describe convolutional, concatenated and turbo code performance, while the maximum correctable bit errors metric is used primarily for block codes.

Convolutional Codes

Most current commercial and scientific digital communications systems rely on convolutional codes to improve performance, including the GSM, USDigital, IS-95 CDMA and UMTS cellular telephone systems, NASA spacecraft telemetry, satellite communications, microwave transmission systems, digital broadcast systems and wireline modem systems. Convolutional

[43] Used in the UMTS ("3G") cellular standard and many current video broadcast systems (set top boxes), among others.
[44] Complexity refers to design and implementation, not computational requirements. In fact compared to Viterbi decoding of low rate (r=1/4) long constraint length (k on the order of 15), turbo codes provide better coding gain with lower computation requirements. [Gumas98]
[45] Note that "coding gain" is normalized to Eb/N0; to determine the actual performance increase if the data throughput is maintained requires accounting for the Eb/N0 decrease if the total carrier power and modulation encoding is not changed and the bit rate is increased to provide the same throughput. (E.g. with a rate ½ code, to maintain 1200 bps throughput necessitates increasing the bit rate to 2400 bps, decreasing the transmitted Eb/N0 by 3 dB; therefore to improve channel performance, the coding gain must be greater than 3 dB.)

coding is implemented at the bit level: Each data bit is transformed into multiple transmitted bits and additionally each data bit is “spread” across many transmitted bits. The number of bits each bit is transformed into is the code “rate” (r); a rate ½ code transforms each data bit into 2 transmitted bits. The number of data bits across which a single bit is spread is the code “constraint length” (k); in a constraint length 5 code each data bit affects 4 subsequent data bits. [Sklar01]

It is simple to implement convolutional coding in hardware or software. Figure 9 shows the GSM standard voice channel convolutional coder, a rate ½ constraint length 5 code, taken from [Hend02]. The information bits are shifted into a 4-stage shift register (labeled “z-1”) clocked at ½ the transmitted bps rate. For each clock cycle two transmitted bits, “G0” and “G1”, are formed by XOR-ing the inputs defined by the connections to the shift register taps. In Figure 9, each “+” is a binary XOR function; the taps are mathematically defined by two polynomials G0(x) and G1(x).

Convolutional decoding can be performed either of two ways.

1. A “maximum likelihood” decoder is typically implemented using the Viterbi algorithm. It evaluates each potential path through the decoding trellis in parallel, so it is time deterministic. This means its decoding time is independent of the incoming bit error rate.

2. A “sequential” decoder is typically implemented with the Fano or a stack algorithm. Sequential decoders are time indeterminate: They follow a path through the trellis but must abandon the current path and try a different one if the metrics deteriorate below a threshold on the current path. For low bit error rates the decoding time is significantly less than a Viterbi decoder, but with a high error rate input bit stream, sequential decoding time rapidly increases to the point where it never completes.

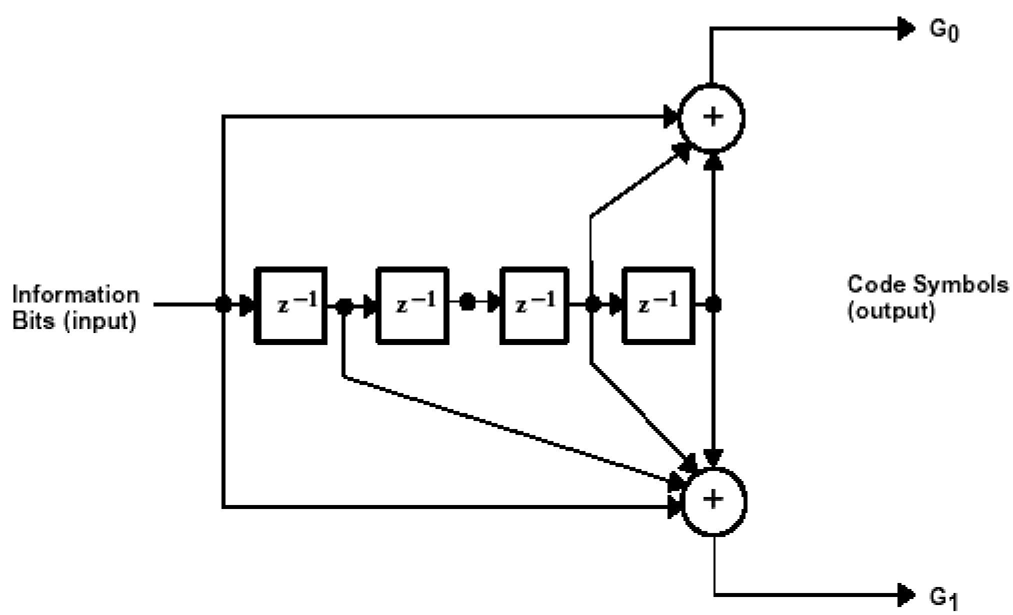


Figure 9: Convolutional coding schematic

Because it is a “maximum likelihood” process, a Viterbi decoder provides higher coding gain than most sequential algorithms in low Eb/N0 environments. A Viterbi decoder can also be more easily implemented in silicon, so for these reasons it is the preferred decoder for high-speed real time communications systems.

Viterbi decoder computational requirements typically increase exponentially as $k * 2^k$, more than doubling for each single bit increase in the constraint length k. [Berro96] This limits current practical Viterbi decoders to a constraint length of 9 or less. The Consultative Committee for Space Data Systems (CCSDS[46]) currently recommends a Viterbi decoder using a constraint length of 7 for its telemetry channel standard. [CCSDS01] A decoder for a constraint length of 7 requires 5.6 times more computation than one for a constraint length of 5.

Constraint length and rate

Convolution code effectiveness increases with longer constraint length. Research for the Galileo planetary mission, summarized in Table 4, shows one example of the increased coding gain with increased constraint lengths. The table is based on a rate $^1/_4$ code to achieve a BER of 10^{-6} using BPSK with coherent detection in an AWGN channel. [Dolin88]

Constraint Length	Coding Gain (To Shannon's Limit, dB)	Increased Coding Gain (k=7, dB)
7	6.4	
9	5.7	0.7
11	5.3	1.1
13	4.9	1.5

Table 4: Convolutional coding gains

In simulation, for a rate ½ code, the Eb/N0 versus BER rate for different low value constraint lengths is shown in Figure 10, taken from [Flemi01]:

[46] CCSDS is an international organization responsible for developing standards for space telemetry enabling interoperability between member space organizations. NASA is a US member agency.

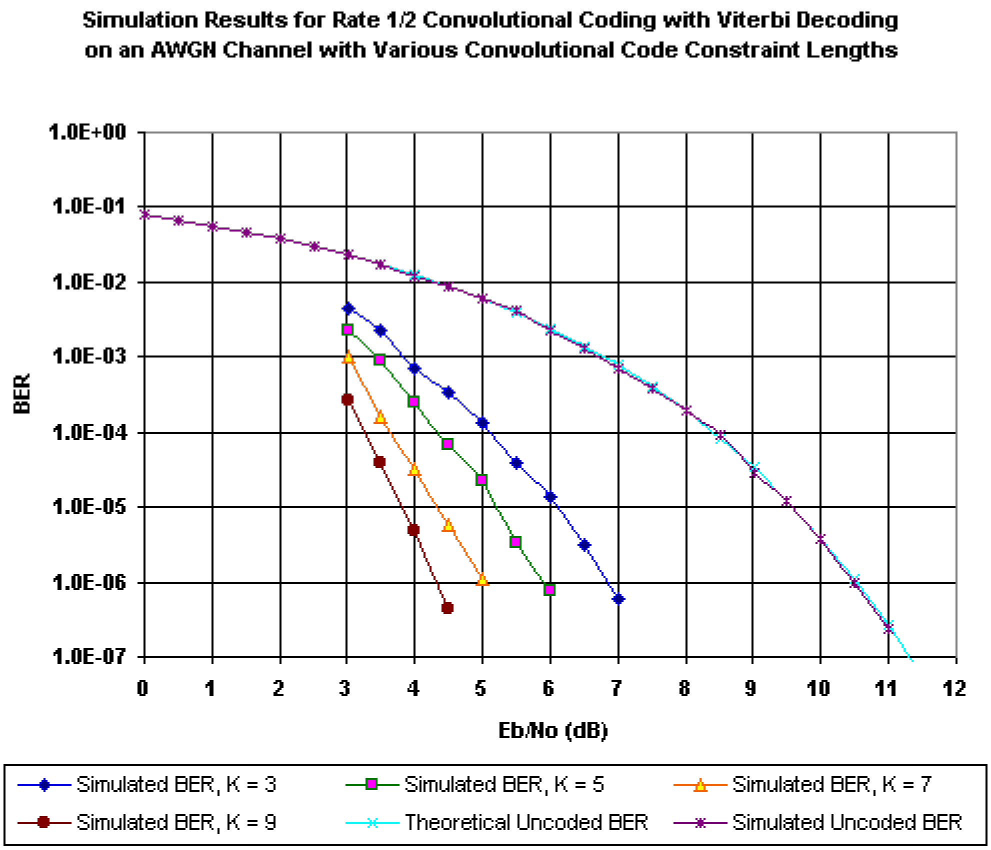


Figure 10: Coding gain as a function of constraint length

Karn's sequential Fano algorithm convolutional decoder for rate ½, k=32 achieves greater than 6.6 dB coding gain in simulation.[47] It executes at a throughput of 152 Kbps (with an input Eb/N0=3dB) on an Intel 486DX4 running at 100 MHz.[48] [Karn95] In another example of a general-purpose microprocessor used for sequential decoding, the Chevillat multiple stack algorithm was implemented in assembler language in 1980 for a Z80 microprocessor. Using a rate ½ code, it demonstrated decoding rates of 630 bps for k=8 and 330 bps for k=15 at an input Eb/N0 of 5.5 dB.[49] [Ma80] Although these decoding rates are an order of magnitude lower than needed for the new ALERT air interface, today's low power, low cost, 16 bit RISC architecture embedded microprocessor achieves significantly more than 10 times the Z80 performance, so it appears feasible.

Related convolutional coding techniques

Most demodulators output a signal as either a "0" or a "1", and the convolutional decoder is forced to use "hard-decision" decoding. Convolutional decoding performance is improved if the demodulator quantifies each received bit value with greater resolution than just "1" or "0". The decoder can use this additional information in "soft decision" decoding to more accurately calculate the trellis path metrics. With a resolution of three bits, coding gain is typically improved by 2 dB.

---

[47] It achieves a BER 1×10^{-5} at 3.3 dB Eb/N0 for BPSK & QPSK in an AWGN channel. This decoder is written in C and available for free under the terms of the GNU Public Software License.

[48] Although it's not clear whether this is achieved with soft decision or hard decision, the software supports both; most likely the coding gain is quoted using soft decision.

[49] RAM requirements were 18 and 45 Kbytes at k=8 and k=15 respectively

In addition to constraint length and code rate, the performance of a convolutional code is highly dependent on the appropriate selection of the generator polynomials and to some degree on matching the code rate to the type of signaling. For non-coherent orthogonal signaling a code rate of ½ is optimum; using $^1/_3$, $^2/_3$ and $^3/_4$ codes results in a coding gain degradation of 0.3, 0.5 and 0.3 respectively. [Sklar01]

Another useful technique is called "punctured codes". In this method, not all the convolutionally coded bits are transmitted. For example, if the data bit stream is encoded at rate ½, but only 3 out of every 4 encoded bits are transmitted, the effective code rate is increased to $^2/_3$: for every 2 data bits, 3 coded bits are transmitted. Punctured convolutional codes sacrifice a small amount of coding gain, yet provide a significant increase in effective data rate. A typical loss is only 0.7 dB when moving to rate $^2/_3$ from rate ½, where k = 7 with BPSK modulation and a coherent demodulator.[50] [Intel02] This small loss increases the effective bit rate 50%; for example, a new ALERT air interface bit rate of 3600 bps using a rate ½ code provides 1,800 bps effective throughput, while a rate $^2/_3$ increases that to 2400 bps

Convolutional coding summary

The salient characteristics of convolutional coding for the new ALERT protocol are:

1. Improved performance in AWGN channels, therefore applicable to ALERT's power limited situations;

2. Arbitrary data length. With no predefined "data block" length there is less decoding latency;

3. A multiplicity of well defined code standards to select from;

4. Simple implementation. Encoding is straightforward, low cost, with a low power requirement; decoding is much more computationally intensive, but is only required at repeater and base station installations;

5. Numerous available solutions. The methodology and its implementation are available in digital signal processors, microprocessor software, FPGA "logic cores" and dedicated integrated circuits.

A side benefit of using convolutional code is that it randomizes the bit stream. Practical non-coherent demodulators perform best when there is a high rate of bit transitions, and an appropriate convolutional code helps to maintain an even bit density in the bit stream at the modulator. This eliminates the need for a pseudo-randomizer, and minimizes the DC accuracy requirement for the audio input and discriminator output interface to standard FM transceivers.

Convolutional code will be an essential component of the new ALERT protocol if we are to achieve significant gains in the bit rate without increasing packet error rates. A logical choice is the NASA standard of rate $^1/_2$, k=7, which yields a coding gain of more than 5 dB. It provides a reasonable trade-off among computational difficulty, coding gain and potential availability of ASSP implementations.

Knowledge gained about actual bit error rates and computational demands during bench testing of the new modulation and Viterbi decoders may lead us to revise this recommendation. For example, we may want to implement a punctured rate $^2/_3$ code, and it could be desirable to increase the constraint length to 9 or reduce it to 5, depending on the tradeoffs between required gains and minimizing complexity.

[50] With the punctured code decoder using a traceback of 70 bit pairs versus 38 bit pairs for the non-punctured code.

Reed-Solomon Code – Block Codes and Burst Noise Protection

Block codes add a certain number of bits to a predefined block of data. These bits carry information about the block that allows the decoder to detect errors, or in the case of FEC codes, actually correct them. Among the simplest block codes is the ubiquitous "even" or "odd" single bit parity added to an ASCII encoded RS232 serial byte. Another is the cyclic redundancy check (CRC) used in such protocols as the "Xmodem" file transfer protocol. Both of these methods detect a large percentage of errors, but are not used to correct them.

The Reed-Solomon code is an FEC block code that offers a high degree of correction for a relatively small overhead. It is particularly well suited for correcting blocks of contiguous bit errors, such as those generated by burst noise. Reed-Solomon (RS) codes operate on "symbols", which are an aggregation of "m" bits. The symbol length is one of the configuration options used to optimize RS performance for a particular application.

During encoding, "parity" symbols are added to "k" information symbols to produce a "code block" of "n" symbols. The resultant code is called an (n,k) code. It has a length of (m * n) bits, (n - k) parity symbols, and (k * m) information bits. All RS code blocks are of a length $n = 2^m - 1$.[51] To correct "e" errors requires (2 * e) parity symbols; therefore the code nomenclature is just (n, n-2e).

The ability of a block code to correct errors is a function of its mathematically defined Hamming distance. Different families of codes exist with varying degrees of error correction capacity. [Rappa96] [Lin83] Frequently used codes for communication channels include the binary Golay, Hadamard and (typically binary) Hamming codes. Reed-Solomon codes are a class of non-binary Bose-Chaudhuri-Hocquenghem (BCH) cyclic codes that exhibit a high versatility in encoding rates, a high speed of encoding and decoding and significant coding gains. They also achieve the largest possible Hamming distance of any linear code. [Rappa96]

Reed-Solomon codes are based on finite[52] field theory called a Galois field. To correct e errors requires (2 * e) parity symbols; for example, to correct three symbol errors over an RS code block requires six parity symbols. Most interestingly, this error correcting property applies to symbol errors, regardless whether one or all the bits in a symbol are in error. This property affords RS codes their good burst error correction capability. For example, for a symbol length of 8 bits (one byte, m = 8), the Galois field is 255 bytes, and to correct 8 symbol errors requires 16 parity symbols or 16 bytes. The result is a (255, 239) RS code. Note that this code will correct up to 8 symbols in error, which could be as many as *64 sequential bit errors*.[53] RS codes perform their worst for random errors; the example code only corrects 8 single bit errors if they are distributed as one bit error in each of 8 bytes.

The (255,239) RS code requires only 6.6% overhead, yet can correct up to 3.3% sequential bits in error. In fact, this is NASA's standard RS code: it is specified as either RS(255,239) to correct up to 8 bytes in error or RS(255,223) which corrects up to 16 byte errors. [CCSDS01]

The fundamental definition of a specific Reed-Solomon code is by its "Field Generator Polynomial" and the "Code Generator Polynomial." Encoding and decoding requires Galois field arithmetic operations. Unfortunately these operations are not directly available on any general-purpose microprocessor and must be emulated. For example, to emulate a Galois field multiply operation[54] requires a test for 0, two log table look-ups, a modulo add and an anti-log table look-

---

[51] The code block can be extended to $n = 2^m$ or $n = 2^m + 1$. Note that generally Galois fields are defined for $n = p^m - 1$, where p is any prime number.
[52] All arithmetic within a Galois field results in an element within the field.
[53] Assuming the first bit error occurred exactly on a symbol boundary, otherwise the limit is 56 to 63 sequential bits in error.
[54] According to one implementation method.

up. The operations require significant computing power, and the decoding algorithm requires a series of steps, typically using the defined algorithms illustrated in Figure 11. (from 4i2i)[55]

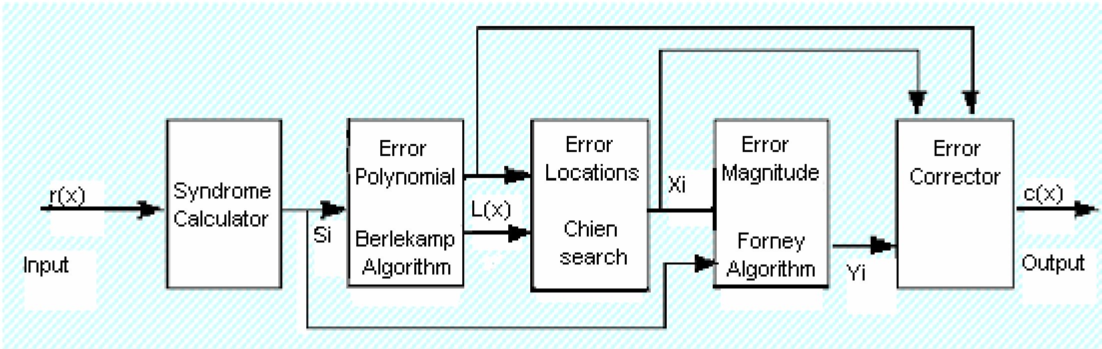


Figure 11: Reed-Solomon decoding

The required computational operations increase as the number of parity symbols increases. Table 5 is a benchmark of the maximum data rate using a commercial software Reed-Solomon decoder that clearly shows this phenomenon. The values are from a 166 MHz Pentium PC executing a commercial software package that is highly optimized (from 4i2i).[56]

Code	Data rate
RS (255,251)	12 Mbps
RS (255,239)	2.7 Mbps
RS (255,223)	1.1 Mbps

Table 5: Reed-Solomon decoding performance

Reed-Solomon code blocks may be arbitrarily "shortened" by transmitting fewer information symbols, yet keeping all the parity symbols. The actual encoding and decoding processes remain identical except that the encoder and decoder both zero-fill all "non-transmitted" information bytes. This technique enables the new ALERT protocol to use the identical code generator and decoder to work with different size messages.

For example, we could choose an 8-bit symbol length and use 4 parity symbols. A non-shortened code block would be 255 bytes, but it can be used to send a short message of 8 bytes of data with the 4 parity bytes, or RS(12,8). The generator would actually pad the 8 bytes of data with 243 "0" bytes in order to generate the 4 parity bytes, then transmit only the 12 meaningful bytes. The FEC overhead increases to 50%, but the error correction capability also increases from 0.8% to 17%.

The identical generator and decoder could be used on another message length, for example a 64-byte block. It would operate as RS(68,64), capable of correcting any 2 bytes in error with an overhead of 6.25%

We conclude that Reed-Solomon is a desirable element of the new ALERT protocol. We believe that implementation is feasible and practical based on the following:

1. Encoding is much less computationally intensive than decoding;

[55] From "Reed-Solomon Codes, An Introduction to Reed-Solomon codes: principles, architecture and implementation", 4i2i Communications Ltd., available at http://www.4i2i.com/reed_solomon_codes.htm as of October 2002.
[56] Ibid. From the same 4i2i web page.

2. Encoding is not a "real-time" task. It can be performed prior to transmission, thus keeping the ALERT gages simple;

3. Although decoding is computationally intensive, it is reasonable to expect a modern high performance microcontroller to perform RS decoding in real time at the anticipated new ALERT rates. Because the convolutional code is likely to be at rate ½, the RS decoding rate will need to be only half the over-the-air bit rate;

4. There is open source code available, in C, implementing a general purpose CCSDS RS encoder and decoder that may be adapted to microprocessor use (available from [Karn02] or [Schul98]).

Concatenated Codes

The use of a concatenated code where an "inner" convolutional code is teamed with an "outer" Reed-Solomon code is synergistic.[57] The convolutional decoder performs best against AWGN, but when it fails, its output error pattern is a burst of sequential incorrect bits. This appears at the input of the Reed-Solomon decoder as a "burst" error, which RS codes excel at correcting. Reed-Solomon codes frequently are used to "clean up" a convolutional decoder output. The text below and Figure 12 are from [Wang02].

> "In 1974, Joseph Odenwalder combined these two coding techniques to form a concatenated code. In this arrangement, the encoder linked together an algebraic code followed by a convolutional code. The decoder, a mirror image of the encoding operation, consisted of a convolutional decoder followed by an algebraic decoder. Thus, any bursty errors resulting from the convolutional decoder could be effectively corrected by the algebraic decoder. Performance was further enhanced by using an interleaver between the two encoding stages to mitigate any bursts that might be too long for the algebraic decoder to handle. This particular structure demonstrated significant improvement over previous coding systems…"

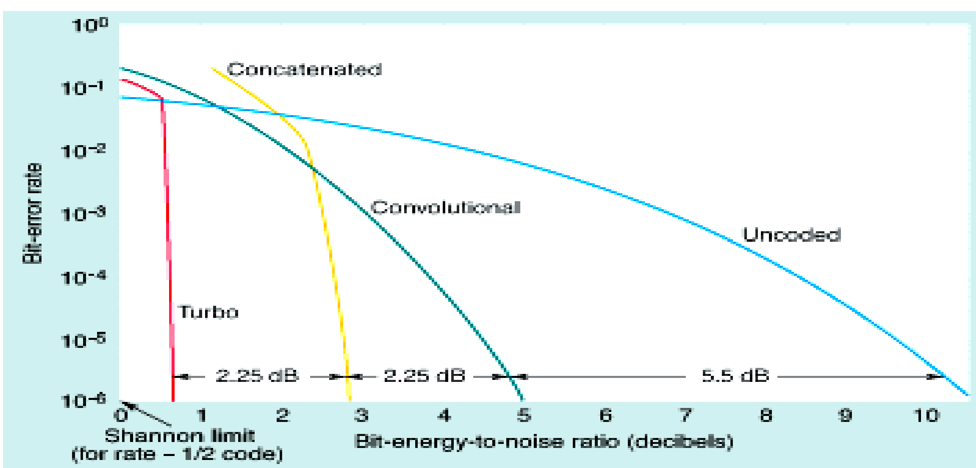


Figure 12: Concatenated coding gains

[57] A general definition for a concatenated code is one that sequentially applies any number of multiple channel codes of any type. Specifically, a highly successful concatenated code is the cross-interleaved multiple Reed-Solomon code used in every CD player.

The NASA standard concatenated code using an inner convolutional code of rate $^1/_2$, k=7 concatenated with a Reed-Solomon (255,223) block code achieves a coding gain of 8.01 dB. Increasing the capability (and computational requirements) of the individual constituent codes increases the overall performance but with diminishing returns. The best code found for the Galileo probe was limited by the encoder hardware, and was a concatenated code with an inner convolutional code of rate 1/6, k=15 with an outer 10 bit symbol Reed-Solomon code. It theoretically achieves 10.12 dB coding gain, which is an improvement of a little more than 2 dB over the NASA standard, but with a massive increase in computational requirements. [Dolin88]

In addition to their synergistic behavior, the sequential architecture of concatenated codes makes them particularly attractive and feasible for the new ALERT protocol. For a real-time decoding of multiple blocks of data, three concurrent operations must take place. While the Viterbi is decoding the bit stream, a second task accumulates and de-interleaves the bits into a block, while a third concurrent task is Reed-Solomon decoding the prior block. Sequential processing allows these tasks to be performed concurrently but independently. One feasible implementation is to partition the decoding tasks among separate small, embedded high performance microprocessors. The diagram in Figure 13 of the NASA standard concatenated code, from [Berle87], is a symbolic representation of the coding/decoding process, but it may also be an implementation architecture:

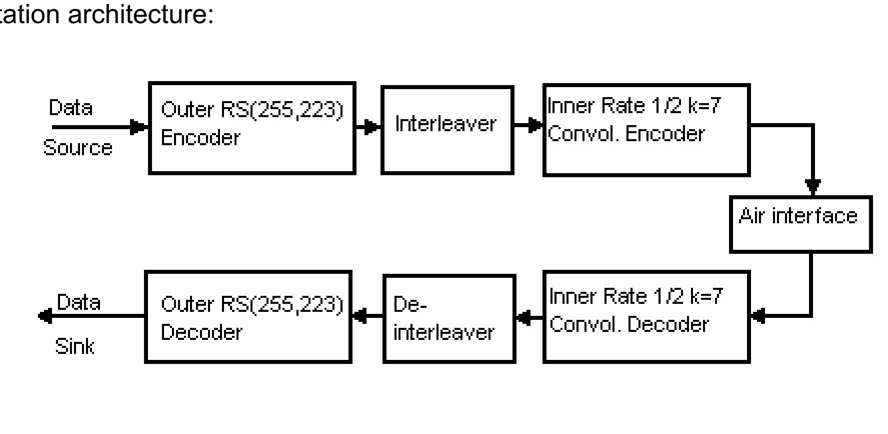


Figure 13: Coding - decoding sequence

Error Detection – CRC or no CRC

The ubiquitous CRC, used predominantly for error detection in packet communications, has a maximum probability of undetected error of 2^{-N} for a CRC of length N bits (assuming the message length is greater than 4*N). [Euro94] The probability of an undetected error using various length CRCs is summarized in Table 6.

CRC length (bits)	Probability of Undetected Error
8	3.9E-03
16	1.5E-05
24	6.0E-08
32	2.3E-10

Table 6: CRC error detection performance

A Reed-Solomon block code not only provides error correction, it also detects uncorrectable errors and exhibits a very low mis-decode rate. Typically, its mis-decode error rate is 5 orders of magnitude lower than its overall BER correction rate. [Berle87] Due to the Reed-Solomon's low undetectable error rate, the Telemetry Channel Coding Blue Book [CCSDS01] standard is to require a CRC error detection only if the channel coding is solely a convolution code. Since the new ALERT protocol includes a significantly shortened Reed-Solomon block FEC code, the probability of an undetected error is orders of magnitude less than using a CRC-16 alone, about equal to a 32-bit CRC. It's unnecessary.

Multipath in ALERT Systems

The second channel impairment that could potentially limit the bit rate of a new ALERT air interface is multipath. A receiver receives the combined sum of all possible direct and indirect signal paths from the transmitter. Since the time of transit for a signal depends on its path length, the signals from different paths arrive at slightly different times. Researchers normally divide this phenomenon into two types:

- Large scale fading is caused by prominent terrain features that reflect, diffract or scatter the signal. Large scale fading is typically modeled as a log-normal distribution function centered about the mean-path loss based on distance.

- Small scale fading refers to the significant changes in amplitude and phase associated with relative motion of the transmitter and receiver, or of reflecting surfaces.

Various statistical models are used to characterize the received signal in a fading channel. The two most commonly used are the Rayleigh and Ricean fading models. The Rayleigh channel is most applicable when there is no line of sight signal path between the transmitter and receiver, where the scatterers are randomly moving, resulting in a zero-mean Gaussian process and where the RF envelope has a Rayleigh distribution over time. The Ricean channel is most applicable when there is a line of sight signal path and/or fixed scatterers and reflectors in addition to the randomly moving scatterers, resulting in a non-zero mean distribution and where the RF envelope has a Rice distribution over time. [Sklar01] [Proak01]

The impact of Rayleigh and Ricean envelope variations can be significant in land mobile channels, but is virtually independent of the symbol rate. The envelope distortions affect only the signal-to-noise ratio and hence Eb/N0, and may require excess path link margins of 10-20 dB to achieve reliable paths. Figure 14 (from [Korn91]) shows the dramatic effect of the channel characterization on the BER due to fading, shown as the impact on the BER for BT = 0.5, GMSK modulation, non-coherently demodulated with a limiter discriminator. The 5 primary curves represent various channel characterizations, defined by K, the ratio (in dB) of the power in the direct path signal versus the non-direct path components. When K = - infinity (the top curve), where only non-direct path power exists, the channel is a Raleigh channel with a uniformly distributed phase. When K = + infinity (curve closest to the vertical axis), there is only direct path power, and the channel is the well-known Gaussian channel (AWGN). The other curves, where K = 16, 10 and 6, represent various intermediate types of channels, all falling into the class of Ricean channels. (The split in the K = 6 and 10 curves at high Eb/N0 represent different demodulation techniques, and can be ignored for this discussion). Note that this fade effect is independent of bit rate – the fading mechanism is dependent on destructive envelope interference from the reflected and refracted signal paths. The effects of this type of fading are already built into the path margins of existing ALERT sites. The addition of FEC to the new ALERT protocol may reduce these requirements.

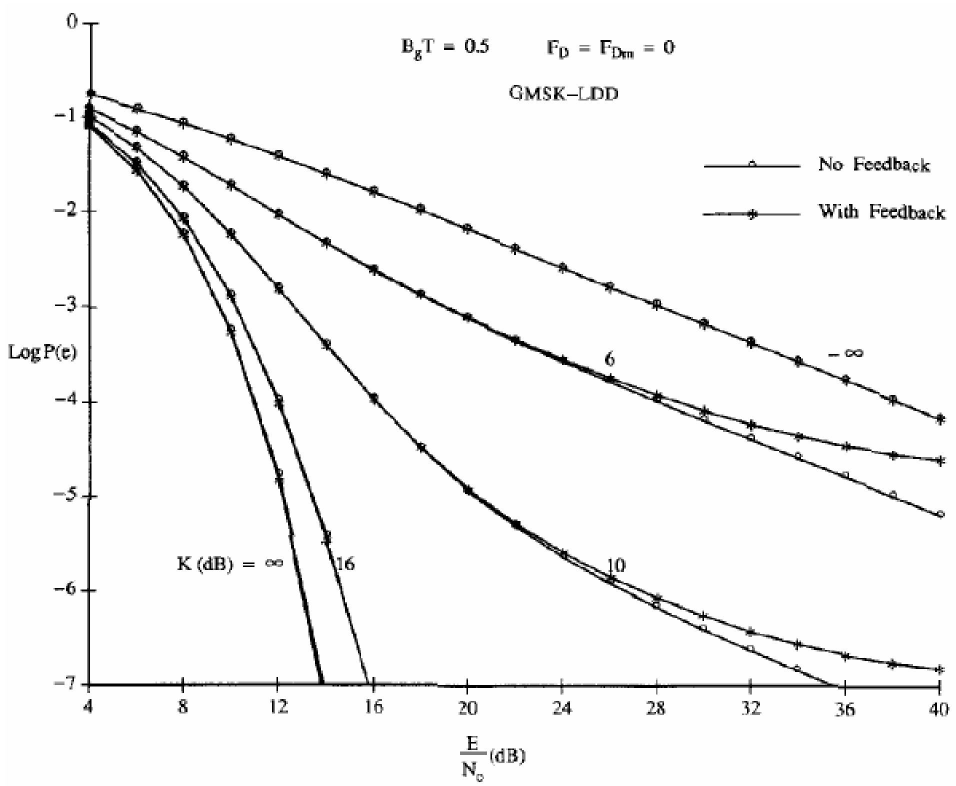


Figure 14: Changes in Eb/N0 vs. BER Due to Rayleigh Channel Fading

Data rate dependent multipath effects

The multipath effect that is bit rate dependent, and therefore concerns us here, is "delay spread." When an RF pulse is transmitted into a channel the difference in arrival time between the first energy pulse and other energy pulses is called the "excess delay." The plot of the received power versus time, based on the transmission of a single pulse, is called the channel's multipath power delay profile (PDF). [58] The maximum time difference from the first energy received to last energy received from a multipath channel is called the channel's "maximum excess delay" (T_M).[59] A Fourier transform can be used to view the time-domain PDF as a frequency domain "coherence bandwidth". If the multipath in a channel is such that $T_M > T_S$ (the symbol period) the channel is classified as a "frequency selective fading" channel, otherwise the channel is classified as a "frequency non-selective" or, more commonly, a "flat-fading" channel.

Frequency selective channels exhibit severe ISI, yet equalization receivers can effectively mitigate the ISI in these types of channels. In flat-fading channels the degradation is primarily the result of destructive signal summation resulting in an envelope fade. [Sklar01] [Chuan87] For discussion purposes, assuming the new ALERT protocol symbol rate is 4800 symbols/sec then its symbol period is 208 microseconds (µsec). Therefore, provided a channel's maximum excess delay is less than 208 µsec the channel is a flat-faded channel. Note that for a delay spread to be 208 µsec the signal path distance difference would need to be 62.4 kilometers (38.8 miles)[60].

[58] Observe that the first pulse received is not necessarily the greatest power; the shorter path may be weaker due to attenuation by obstacles, or if no direct path exists, may be a weaker refracted or reflected signal.
[59] The typical cutoff threshold in measurements is -20 dB.
[60] The maximum selectable multipath delay is 186 µsec on the HP (Agilent) Model 11759C RF Channel Simulator.

Every transmitter and receiver pair exhibits its own unique channel maximum excess delay. Figure 15 (from [Fleur96]) is a pictorial representation of the source for excess delay (shown for a moving receiver). Diagram (a) is the physical environment; (b) is the direction-delay-spread response; and (c) is the multipath power delay profile.

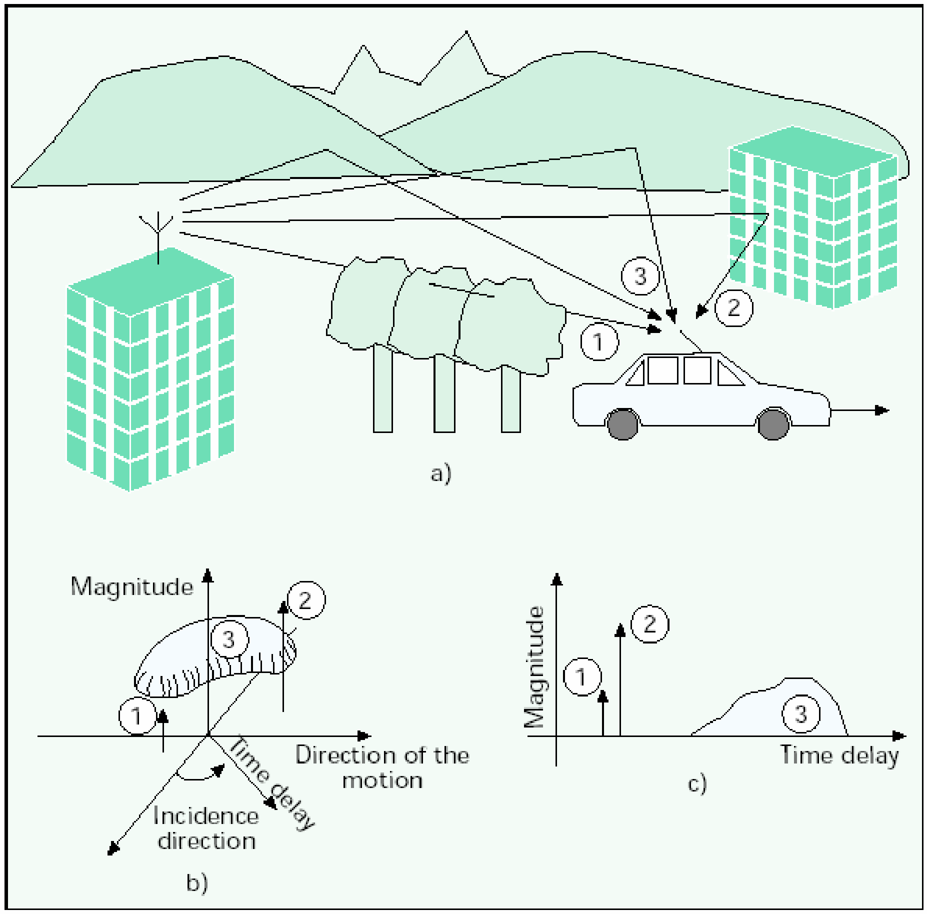


Figure 15: Physical Causes for Multipath

Based on empirical measurements, a number of multipath PSD models are currently used to simulate channel performance for theoretical evaluation and testing. Figure 16 shows the maximum severity European GSM cellular phone multipath model. An impulse transmission creates three sets of exponentially decaying received power peaks: the first at 0-2 μsec with amplitude of 0.33, a second at 30-42 μsec with amplitude of 0.5 and the third at the maximum excess delay of 80-85 μsec with amplitude of 0.17. [Rohd01]

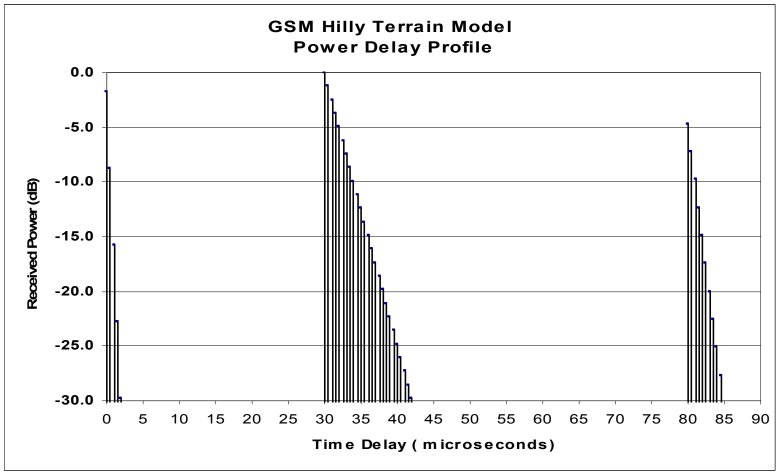


Figure 16: GSM Hilly Terrain Model Power Delay Profile

The COST 207 (European Cooperation in the field of Scientific and Technical Research) "hilly terrain" (HT) model, another 900 MHz severe delay model, is shown in Figure 17. Its maximum excess delay peak is at 19.9 µsec with an amplitude 6.7 dB down from the peak at 1.9 µsec (from [Lee99]).

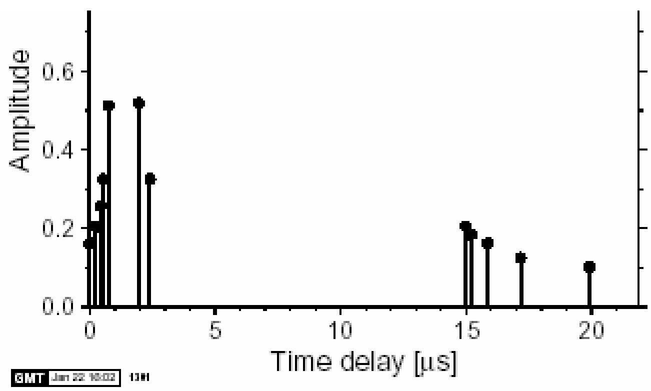
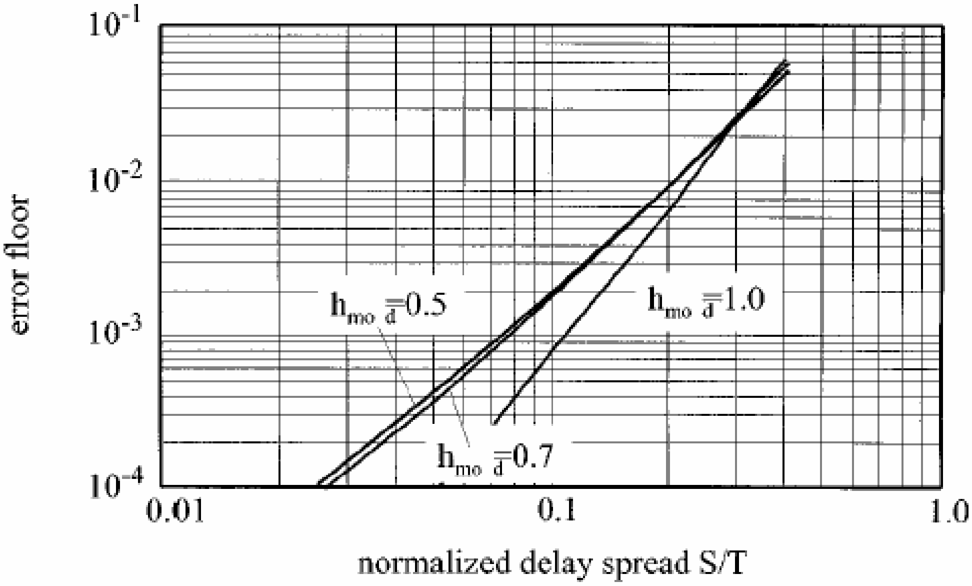


Figure 17: Modeled multipath delays

A channel's excess delay is one factor[61] in what is termed the "error floor" or "irreducible BER" of a channel, the phenomenon where increased power does not improve the BER. For FSK modulation with limiter discriminator detection, analysis has shown that modulation index also has a pronounced impact on the irreducible BER. Modulation indexes of 0.7 and 1.0 have an improved tolerance to excess delay compared to a modulation index of 0.5. The predicted irreducible error rate for binary FSK versus a normalized "RMS delay spread" (T_{RMS})[62] divided by symbol period, T_S is shown in Figure 18 (shown for stationary channels exhibiting no Doppler shift). [Petro00]

[61] Another factor is the Doppler shift of a signal.
[62] Root Mean Square (RMS) of the excess delay power delay profile (PDF). See [Proak01] page 803 for a mathematical definition of T_{RMS}.



T_{RMS} / T_S

Figure 18: Multipath effects on error floor

The chart shows that when the RMS delay spread is 20% of the symbol rate, the error floor is 0.6% BER for FSK with modulation index of 1.0 and approximately twice that for a modulation index of 0.5. The chart also illustrates that when a channel's RMS delay spread is less than 10% of the symbol period, the bit error rate floor is below 0.08% (8×10^{-4}) for a modulation index of 1.

A different view of the irreducible error rate caused by excess delay is shown in Figure 19. It shows the BER rate versus Eb/N0 for three different normalized T_{RMS}/Ts: at 7.4%, 20.5% and 59.5%. [Finton]

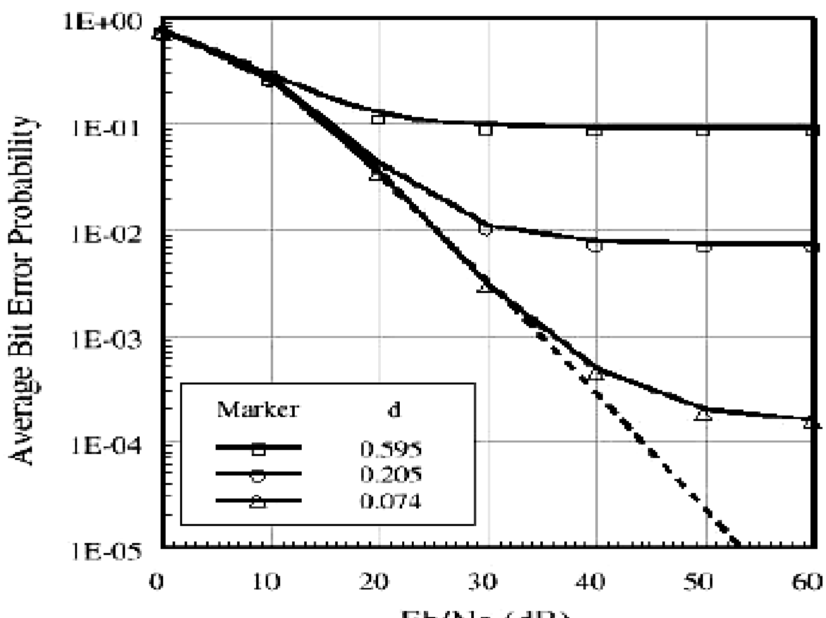


Figure 19: Error rate at various multipath delays

These studies and others by Chuang [Chaun87] confirm that a good rule-of-thumb is:

> "If a digital signal has a symbol duration which is more than ten times the RMS delay spread, then an equalizer is not required for bit error rates better than 10^{-3}. For such low values of spread, the shape of the delay profile is not important." (From [Ander95])

The authors were unable to uncover any theoretical or empirical measurements for the impact of multipath on AFSK-FM data telemetry. In our judgment, multipath BER degradation for AFSK-FM is at best the same as digital FSK, but due to its requirement for a second FM-AM discriminator, AFSK-FM could be much more susceptible to multipath effects.

Mitigating multipath effects

The simplest method of mitigating error rates induced by delay spread is to constrain the new protocol to a symbol period at least 10 times longer than the expected RMS delay spread.

A key concept when applying this rule of thumb, and interpreting the graphs above, is that they all are based on the "RMS delay spread", the *root mean square* of the excess delay power distribution profile, *not* the maximum excess delay. The RMS delay spread, T_{RMS} can be best described as the standard deviation centered on the weighted mean of the power delay profile. Given its typical log-normal distribution this is *significantly* less than the maximum excess delay. Intuitively, since the free-space signal power decreases at a rate proportional to distance squared, delays on the order of 10s of µsec from the initial peak are significantly attenuated and thus contribute only minimally to increasing the standard deviation.

The worst case RMS delay spread model found in the literature, the GMS "hilly model" shown earlier, has an RMS delay spread of 21.7 µsec even with its maximum excess delay (using a 20 dB threshold) of 83 µsec. A more typical "hilly" terrain model, the COST 207, has an RMS delay spread of 6.8 µsec and a maximum excess delay of 19.9 µsec.

The 22 µsec RMS delay spread of the worst case model is only 7.9% of a 3600 symbols/sec symbol period and 10.6% of a 4800 symbols/sec symbol period. Most RMS delay spreads seen in the field will be much less than 22 µsec and it appears feasible to increase the symbol rate up to 4800 symbols/sec and maintain a reliable communications path.

In fact, although "[maximum] Excess delays as high as 100 µsec…can occur in practice…" [Petro00], a number of studies verify that an RMS delay spread of 22 µsec is severe even for the mountainous geography frequently encountered in ALERT systems. The "average extreme case" for a suburban environment (using a 10 dB threshold) measured an RMS delay spread value of 1.96 to 2.11 µsec. [Rappa96]. The most severe terrain-based SUI[63] model (with high delay spread and high Doppler shift), developed from empirical measurements across the continental US for simulation of MMDS (900 MHz) propagation, has a maximum excess delay of 20 µsec and an RMS delay spread of 5.2 µsec.

In 1988 the Swiss Post, Telephone and Telegraph, General Radio Technology Section studied multipath power delay profiles for 8 receive sites in the Alps around Lake Geneva and Lake Lucerne. [Weck88]. Their reported worst case RMS delay spreads, as reported by the 90th percentile, were significantly less than 22µsec. By using a mobile pulse transmitter broadcasting at 900 MHz, the study recorded over 70,000 PDFs in support of defining propagation

[63] The SUI-6 model of the Stanford University Interim (SUI) Channels Models. From "Channel Models for Fixed Wireless Applications," January 22, 2001. See 802.16 working papers at: http://www.ieee802.org/16/tg3/contrib/802163c-01_29.pdf

characteristics for the GSM cellular system. The key propagation parameters were determined from each PDF and statistically reported for discrete percentile ranges for each of the 8 sites. The 90th percentile RMS delay spread was 10.0 µsec at 2 sites; another site was 9.1 µsec and all 5 other sites were below 8 µsec. The 90^{th} percentile for maximum excess delay was 23.8µsec across all sites.

Figure 20 shows two representative multipath delay profiles from a recent study of the multipath in mountainous terrain, taken near Cowichan Bay, close to Victoria, BC, Canada (taken from [Dries00]). The study measured the delay profiles from four separate sites under various conditions. The maximum excess delay measured from any delay profile was 31.25 µsec. The distribution of measured maximum excess delays from all PDFs was: 11% greater than 20 µsec, 46% in the 10-20 µsec range and the remaining 43% less than 10 µsec.

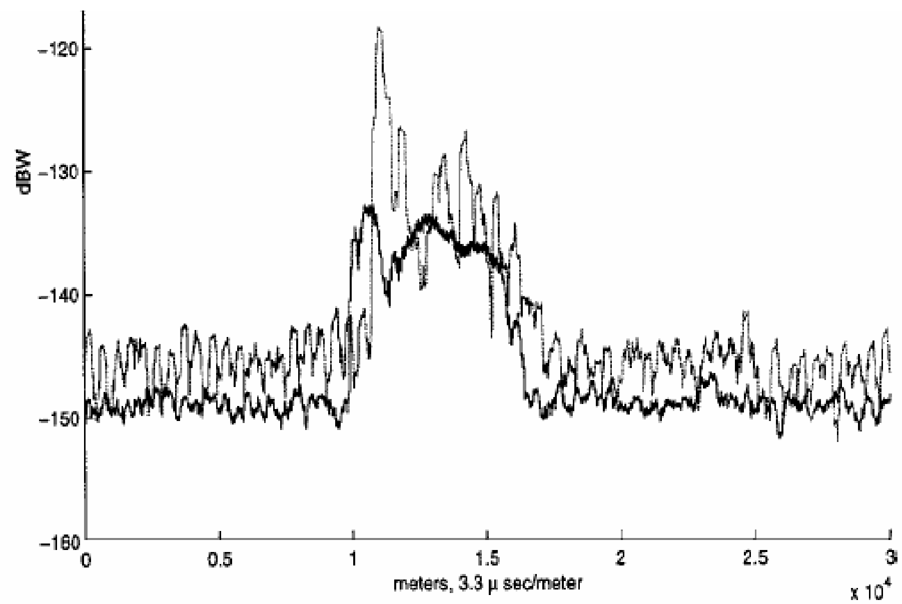


Figure 20: Sample measured multipath delay profile

Fortunately, the presence of forward error correction in the proposed new ALERT protocol significantly mitigates multipath effects. The meaning of a multipath induced "irreducible" bit error rate means "not reducible by increasing EIRP." The delay spread error mechanism is still a random process that only distorts a percentage of received bits. Since the FEC process spreads individual bit information across multiple bits, it error corrects the received bit stream whether the error is caused by random noise or ISI due to multipath. In fact, "Error-correction coding, coupled with interleaving, is probably the most prevalent of the mitigation schemes used to provide improved system performance in a fading environment" (from [Sklar01] page 984).

To further avoid the effects of multipath, it may be prudent to use four-level filtered FSK to achieve a bit rate of 4800 bps at a coded rate of 2400 symbols/sec. This doubles the symbol period to 417 µsec, which (from Figure 18) reduces the worst case GSM model RMS delay spread to 5.3% of the symbol period, and improves the irreducible BER to better than 10^{-4}.

Whether GMSK or GFSK modulation is feasible or suitable in multipath environments is summarized in Table 7. In evaluating modulation techniques for the 802.11 wireless LAN physical interface, California Microwave presented the following chart rating various modulation techniques against channel impairments, including "delay spread sensitivity." It clearly shows quaternary filtered FSK to be the best, with the other forms of binary filtered FSK ranking second (from [Silbe93]).

Parameter	4CPFSK	2CPFSK	MSK	0.39 GMSK	xPSK	xQPSK	4 QAM
Sensitivity to Envelope Variations	no; limiter used	no; limiter used	no; limiter used	no; limiter used	yes / linear amplifier required	linear amplifier required	linear amplifier required
Sensitivity to Phase Impairments	no	no	yes	?	phase flips phase glitches osc. phase noise	phase flips phase glitches osc. phase noise	-phase flips -phase glitches osc. phase noise
Interference performance	A-	A	B	B-	C	C	C
Suitability to Frequency Hopping	A	A	B	A	?	?	?
Demodulation	non coherent	non coherent	non coherent	non coherent	coherent preferred differential possible/ not suitable	coherent preferred differential possible/ not suitable	coherent preferred differential?
ISI / Delay spread sensitivity	A	B	B	B	C	C	C
Robustness to Co-channel Interference	Excellent	Excellent	Excellent	Good	Good	Good	Poor

Table 7: Characteristics of various modulation techniques

Conclusions about multipath

Rayleigh and Ricean multipath fading affects the communication reliability in a channel regardless of bit rate. The addition of FEC to the new ALERT protocol may mitigate this aspect of multipath fading.

The easiest approach to avoiding data rate-dependent multipath interference is to keep the air interface symbol period more than ten times the RMS delay period. The difficulty is arriving at reliable estimates of RMS delay over the range of actual ALERT terrain conditions and path distances. The RMS delay spread used in the GSM hilly terrain model appears to be pessimistic; available empirical studies suggest that a 22μsec delay is a good working worst-case limit for ALERT environments.

We believe that at 3600 symbols/second multipath-induced BER increases will be negligible. Even at 4800 symbols per second, a 22 μsec RMS delay spread is just 10% of the symbol period, and therefore the irreducible multipath BER will be approximately 10^{-3} even without any FEC. Since forward error correction further mitigates the effects of multipath, we believe that a 4800 symbols/sec rate should be a part of any field trials. To minimize multipath impact, the modulation index should be maximized.

Physical Packet Structure

Field demonstrations and bench tests of the new air interface will be required to finalize a new ALERT protocol specification. While the data content of a new protocol is not the concern of this study, there must be a defined over-the-air packet structure that can be used to test the proposed modulation and channel coding techniques. Current industry best practices, combined with the specific needs of the ALERT community, largely define the shape of the new physical packet structure. It should be straightforward to specify a packet structure that can encompass multiple data formats and support enhanced upper layer protocols.

Modern communications protocol concepts include layering, encapsulation and packetization. Encapsulation enables operational flexibility and structured application development. For instance, by separating the application data from communications dependent information such as start and stop bits, the application does not need to know how the data was communicated. This isolation allows the application to use multiple communication channels without modifying the data passed to the communication handler. At the operational level, if it's not necessary for the communication handler to inspect data content of the packet, the same communication handler, and therefore channel, can be used to transfer a variety of types of data.

The proposed physical packet structure embodies the concept of a packet header that can be used as a stand alone message. When used as the minimum sized message it contains a small data field as well as the fields necessary to interpret the message. When an optional data block follows the header specifies its length and format. The use of a header section followed by variable data blocks is well established in current communications technology.[64]

The physical packet structure we propose is composed of three sections: a preamble, a header and the optional data block.

1. The 144-bit fixed length preamble contains the initial 96-bit synchronization bit stream and the 48-bit frame synchronization pattern bit stream.

2. The 12-byte fixed length header contains a:

    - 16-bit source address,

    - 8-bit protocol identifier field (the PID, describing how to interpret the header data field and optional appended block),

    - 8-bit reserved field,

    - 32-bit data field, and

    - 32-bit Reed- Solomon FEC applied to the header.

3. The optional, variable length data block contains the data and the 32-bit (or more) Reed-Solomon FEC applied to the data block.

An initial carrier transmitted without modulation, as is sent with current ALERT transmissions, is not required with the candidate GMSK or GFSK demodulators using "open squelch" with long bit sync and frame sync patterns.

The header and data block sections are convolutionally coded, increasing their bit size dependent on the convolutional code rate plus the required convolutional code tail.[65]

---

[64] For example TCP, UDP, IP, ATM, Ethernet.
[65] Since encoding spreads each bit across k/r transmitted bits, an extra k/r bits must be transmitted to "flush" the last data bit through the system. Note that for header packets without a data block, the tail must be added to the header; those with an appended data block only require the tail at the end of the data block.

Transmissions containing just the header are short and minimize channel use.[66] Frequently header-only transmissions are used in two-way protocols (a possible future ALERT option) for link coordination such as transmission acknowledgements (ACKs), repeat requests, clear to send, request to send etc. The protocol identifier field in the header identifies to the receiver how to interpret the header packet and whether or not a data block is appended.

We propose to use a header-only transmission in the new ALERT protocol to send simple ALERT messages with the least overhead. Since currently these simple ALERT messages are multicast, there is no requirement for a destination address. The 32-bit data field in the header packet is used to send a 16 bit ALERT sensor ID and a 16-bit sensor data value. Not only does this enhance the ALERT ID and data value to 16 bits each, it adds a source address to the fundamental ALERT protocol. The key is to define a specific PID field value to this type of header-only transmission. (See Appendix B – ALERT Special Header-Only Packet.)

Using a different PID value enables larger data content transmissions by appending the data block. The data field in the header packet is then used to define the length of the appended data block, the communication and FEC encoding of the data block and the destination address for the data content. These are mapped into the data field as:

- A 16-bit destination address;
- An 8-bit data block length field (specifying the data content length in the appended data block, modulo 4 bytes);
- An 8-bit data-encoding field that identifies the type of FEC and/or modulation.

This physical packet structure is flexible. The header identifies the amount of data embedded in any appended data block in increments of four bytes. Using 4-byte resolution and an 8-bit field enables packets to be tailored to their specific data content requirement, from more than 1,000 data bytes to as few as 4 bytes. Since size is determined in real time, gages can adjust the data sent depending on conditions. The protocol does not require all gages to send identical size packets: various types of gages, each transmitting various size packets, can easily share a channel. A 28-byte data packet is described in Appendix B. This example data content size was selected solely for comparative metrics; it is sent in about ½ the time compared to the present ALERT message.

Each PID value specifies a unique data format, enabling more efficient data transmissions. For example, by defining a fixed format for sending a common set of multiple sensors as a unique PID, there is no longer a requirement to send individual IDs for every sensor value in the transmission. Also, including a source address field in the header enables the philosophical split between a sensing location and its one or more sensor IDs.

The preamble itself must be robust enough to provide a solid bit clock lock and exact frame synchronization for FEC decoding in low Eb/N0 environments, using a variety of different demodulators. A 48-bit synchronization word with strong autocorrelation, proposed as a triple concatenation of the NASA standard 16-bit frame sync pattern, is thought to provide strong protection against false locks. The preamble is considered conservative; through bench testing it may be found that smaller values for both frame sync and bit sync are acceptable.

This proposed physical structure will enable us to evaluate the modulation and coding technologies and gain practical experience with the application of modern protocol structures to the ALERT environment. The physical packet structure and the air interface we wish to test are

[66] The preamble is always sent prior to any header block for demodulation; hereinafter, any reference to sending a packet implies sending the preamble prior to the header.

inter-related, but only loosely. Future modifications to this candidate structure will be possible, reasonably independent of the new ALERT modulation and FEC coding methods.

Transition Plan

The new ALERT modulation and coding schema will be inter-operable with the existing 300-bps, four-byte message on the same frequencies. Obviously, the channel capacity will be constrained by the sum of all transmissions, regardless of their type. However, the modulation, data rates and encoding schemes are so different that each will be appear essentially as noise to the other.

The new protocol will be received at the base station as the audio output of the receiver's discriminator. The decoding process will take place in a dedicated modem that handles the convolutional decoding, the de-interleaving, and the Reed-Solomon algorithms. What is delivered to the base station computer is a serial data stream in the format(s) specific to the data block types. In order to make use of this data, the ALERT base station software will need to recognize one or more of the new formats. This amounts to little more than a series of parsers, each triggered by a specific PID value in the header.

The new air interface operates "open squelch", so it is not yet certain whether the same receiver can be tapped to supply audio input to both the old and new decoders simultaneously. The worst case, (which is also probably the simplest and most practical case) is that separate receivers would be used. Most likely, the decoded serial output will be fed to separate ports on the base station computer.

With upgrades to the base station software, receiver and decoder in place, an ALERT user could intermix old and new protocol ALERT equipment in the field. There would be no need to systematically replace all equipment in order to begin using the new protocol.

The new protocol devices will require their own repeater, where one is needed. The repeater will perform the same decoding and message assembly functions as at the base station, then reconstitute and retransmit passed messages. We recommend that the repeater incorporate the ability to receive messages in the old format, and translate these into header-only messages in the new format. This would eliminate the introduction of errors in the old format data at the earliest opportunity. One of the significant technical challenges of the new protocol will be to incorporate the required processing complexity into a low-power, remotely operated repeater package.

Another aspect of the transition plan relates to the manufacturers of ALERT equipment. Existing ALERT equipment that produces only a sensor ID and data value could be interfaced to an OEM terminal network controller (TNC) that encodes the data in a simple implementation of the new protocol. Obviously, the full potential of the new protocol would await the development of more extensive capabilities in ALERT products. Existing loggers and remote terminal devices that are capable of serial output could be easily adapted or configured to make use of the new protocol. As the demand arises, new data block types within the protocol could be devised, managed or sanctioned by the ALERT Users Groups.

Conclusion

We surveyed the current integrated UHF/VHF radio modems for land mobile applications from the perspectives of cost, inter-operability and open systems. From that, we determined that it is not possible to adopt an existing commercial integrated radio modem. The results did validate the concept that 4800 bps or higher is currently routine in a standard 12.5 kHz channel and that forms of filtered (Gaussian or Raised Cosine) FSK or MSK are the modulation technologies of choice.

A review of commercial telemetry protocols determined that no current open wireless protocol appeared suitable for adoption. Most open protocols are not targeted to extremely short multicast messages. Protocols that might be candidates, such as MDC-4800, are not open. The open protocols do not have the robust FEC and frame synchronization needed to maintain the current ALERT radio coverage, and/or use legacy mechanisms such as “bit stuffing” that significantly complicate block-based FEC mechanisms. The review did confirm that Reed Solomon or convolutional forward error correction codes and 4-Level GFSK modulation schemes have been adopted by the newer wireless protocols.

We investigated two potential limitations of moving to a higher bit rate while maintaining the same radio coverage. The first is the increased bit error probability incurred when the energy per bit is decreased by a higher bit rate. The second is the increased potential effect of multipath on a higher bit rate air interface.

The energy per bit decrease is approximately 12 dB when increasing the bit rate from 300 bps to 4800 bps. We showed that this loss can be overcome by a combination of new modulation technologies and channel coding. The present ALERT protocol uses a legacy Audio FSK over FM for which there is a dearth of information that would allow direct performance comparison to any of the potential digital FSK technologies. Our best estimate is that digital FSK offers a 6 dB improvement over Audio FSK at any given data rate. Our recommendation is to use using either binary RC FSK(CPM) or GMSK.

Another 6.6 dB of coding gain can be expected by using a concatenated channel code consisting of an inner code based on NASA’s standard rate $^1/_2$, k=7 convolution code and a significantly shortened Reed-Solomon 8-bit symbol outer block code. In theory, a 4800 bps air interface should exhibit a BER rate similar to that for the 300 bps FSK currently used for ALERT transmissions.

We discussed two additional methods to further increase the bit rate, though with small degradations in BER and increased demodulator complexity. The first is to use a punctured rate $^2/_3$ convolutional code, and the second is to use a four-level RC FSK(CPM).

It was also shown that it is feasible to decrease the BER by using a longer constraint length code, although the Viterbi decoding may be impractical due to the computational requirement. Conversely, the convolutional code constraint length could be reduced to 5, which would result in a combined coding gain of 5.8 dB rather than 6.6 dB. It this configuration proved acceptable, it could be implemented at a more than five-fold reduction in Viterbi decoder processing.

We found that current research validates the “industry rule of thumb” regarding multipath problems with a higher bit rate air interface: As long as the RMS delay spread is 10% or less of the symbol period, the irreducible bit error rate should not be affected by multipath, regardless of whether there is a line of sight path to the transmitter. Research confirmed even in extremely severe multipath environments such as the Canadian Rocky Mountains, that although excess delay may reach 100 μseconds, the RMS delay spread is less than 10% of a 3600 bps symbol period. We concluded that multipath will not impact 3600 symbol per second transmissions, and that significant impacts on a 4800 symbol per second air interface are unlikely.

We showed how the new air interface can be incorporated into a flexible physical structure by using a packet containing a preamble, header and optional data blocks. Minimal sized ALERT messages are accommodated, yet the structure has the ability to transmit over 1,000 bytes of data using a variable packet size with a four-byte granularity.

It is feasible and practical to perform the new modulation with low-cost, low-power microcontrollers. This is demonstrated by the market availability of low cost complete Amateur Radio modems that interface to off-the-shelf radio transceivers using RC FSK (CPM) and GMSK for 9600 bps and higher modulation. Additionally, an ALERT gage modem is not required to execute the demodulation/FEC decoding tasks; its only requirement is to encode. The Reed-Solomon encoding, interleaving, and convolutional coding can all be performed as "pre-transmission" tasks. Therefore the gage computation requirement is small and only memory (RAM) may need to be increased.

ALERT repeater and base station demodulators and decoders will require significantly more computational power, either in general purpose microprocessor and software, DSP hardware and software, FPGAs or application specific integrated circuits. Power and performance requirements are difficult to determine at this juncture, but a conservative estimate based on current research is that an embedded PC card is the upper bound of what would be needed. In the authors' opinion, an architecture that uses multiple low-power, low-cost, 16-bit RISC embedded controllers may be all that is required.

The next step following this study is prototype development, bench testing and real world pilot testing.

> "Although many researchers have been working hard during the past few decades in the area of field strength prediction, there are still numerous problems to be solved. To date a perfect propagation model has not been found. Instead, many different prediction tools have been proposed. Each has advantages and disadvantages and can only be applied to particular circumstances." [Nesko00]

Appendix A: Integrated Radio Modems

UHF (MHz)	VHF (MHz)	Channel BW (KHz)	FCC Emission Designators	Over the Air bps	Modulation	BER	Sensitivity (12 dB SINAD)	Notes
Data Radio COR Ltd.								
T-96SR Modem + Transceiver								
380-512	132-174	12.5	9K30F1D	4800, 9600	DRCMSK:	1x10-5 @ -107 dBm (1.0 uV)	<-116 dBm (0.35 uV)	Differential Raised Cosine Minimum Shift Keying; 7 bit scrambler; Differential (NRZI); deviation +/-2.5 KHz @4800bps; T-96SR has no packet defined; RTS to CTS delay is approx. 30 msec (clear carrier); No FEC
380-512	132-174	25	9K30F1D	4800, 9600	DRCMSK	1x10-6 @ -107 dBm (1.0 uV)	<-116 dBm (0.35 uV)	BER spec. is for 9600 bps, "4800 is better"per discussion with Bob Tnka.
380-512	132-174	25	15K3F1D	19200	DRCMSK	2x10-5 @ -102 dBm (1.7 uV)	<-116 dBm (0.35 uV)	
Integra-TR Modem + Transceiver								
380-512	132-174	6.25	6K00F1D	2400	DRCMSK	1x10-6 @ -116 dBm (0.35 uV)	<-116 dBm (0.35 uV)	Differential Raised Cosine Minimum Shift Keying; 7 bit scrambler; Differential (NRZI); deviation +/-2.5 KHz @4800bps; Header length is 4 msec (includes frame/bit sync); no FEC, CRC-16 error detection; fast RTS/CTS turn around at 4 msec.
380-512	132-174	12.5	9K30F1D	4800, 9600	DRCMSK	1x10-6 @ -104 dBm (1.4 uV)	<-116 dBm (0.35 uV)	
380-512	132-174	25	9K30F1D	9600	DRCMSK	1x10-6 @ -107 dBm (1.0 uV)	<-116 dBm (0.35 uV)	
380-512	132-174	25	15K3F1D	19200	DRCMSK	1x10-6 @ -100 dBm (2.3 uV)	<-116 dBm (0.35 uV)	
Simrex Corp. Aria-GLB								
SNRDS Modem + Transceiver								
218-235, 406-512	150	6.25, 12.5, 25.0		1200-19200	FSK, GMSK			Modem probably uses CMX589 IC
218-235, 406-512	150	25		9600	GMSK	1x10-6 @ -113 dBm	-119dBm	Turn around time <5msec; MX.25 protocol: 256 byte packets; optional FEC (linear block code - Hamming Code), (spec'd BER is w/o FEC ?); proprietary

Appendix A: Integrated Radio Modems

UHF (MHz)	VHF (MHz)	Channel BW (KHz)	FCC Emission Designators	Over the Air bps	Modulation	BER	Sensitivity (12 dB SINAD)	Notes
Vytek Products								
FireLine Modem + Transceiver (OEM)								
450-470		25	16K0F1D	19200	4-Level GFSK	1x10-3 @ -100 dBm		"Octagonal Signaling (pat. pending)"; turn around time <10 msec
450-470		12.5	9K30F1D	9600	4-Level GFSK	1x10-3 @ -104 dBm		"Octagonal Signaling (pat. pending)"; turn around time <10 msec
450-470		25	16K0F1D	9600	2-Level GFSK	1x10-3 @ -106 dBm		"Octagonal Signaling (pat. pending)"; turn around time <10 msec
450-470		12.5	9K30F1D	4800	2-Level GFSK	1x10-3 @ -110 dBm		"Octagonal Signaling (pat. pending)"; turn around time <10 msec
Skyline Modem + Transceiver								
450-470	136-174	25		9600	GMSK	1x10-3 @ -110 dBm		packetized; 16 bit multipoint addressing; CRC-16 error detection
450-470	136-174	12.5		4800	GMSK			packetized; 16 bit multipoint addressing; CRC-16 error detection
Pacific Crest Corporation								
TM32 Modem + Transceiver								
Yes	Yes	25,12.5,6.25		2400, 4800, 9600	GMSK (BT 0.3 or 0.5)		-114 dBm	Hamming Code (12,8) FEC with interleave, up to 20 consecutive bit errors (packet or transparent mode), 50% overhead; synthesized with fast turn on <10 msec
EDL Modem + Transceiver (OEM model PDLRXO has similar specs. $700 ea. in qty. 1-99)								
Yes	No	25		19200	4-Level FSK			Hamming Code (12,8)
Yes	No	25		9600	GMSK	1x10-5 @ -110 dBm		unclear whether this is with or w/o FEC Hamming Code (12,8)
Yes	No	12.5		9600	4-Level FSK			Hamming Code (12,8)
Yes	No	12.5		4800	GMSK			Hamming Code (12,8)
RFM96W Modem + Transceiver (OEM version is RFMRXO)								
Yes	Yes	25,12.5		4800, 9600	GMSK (BT 0.3 or 0.5)		-117 dBm	Hamming Code (12,8) FEC with interleave, up to 20 consecutive bit errors (packet or transparent mode), 50% overhead; synthesized; up to 35W

Appendix A: Integrated Radio Modems

UHF (MHz)	VHF (MHz)	Channel BW (KHz)	FCC Emission Designators	Over the Air bps	Modulation	BER	Sensitivity (12 dB SINAD)	Notes
RF DataTech								
QRT Modem + Transceiver								
406-512	138-175	12.5		1200	FSK, Bell202	1x10-3 @ -120 dBm	-119 dBm	
406-512	138-175	12.5		4800	GMSK	1x10-3 @ -117 dBm	-119 dBm	
406-512	138-175	12.5,20,25		1200-19200	FSK; FFSK(MSK) 2400; GMSK		-119 dBm	Internal modem includes POCSAG paging at 512, 1200 & 2400; Differential encoding (NRZI) selectable on/off; MSK at 1200/2400 or 1200/1800
SRT Modem + Transceiver								
406-512	138-175	12.5		9600	4-Level FSK	1x10-3 @ -115 dBm	-117 dBm for 12 dB SINAD (flat)	
Electronic Systems Technology								
ESTeem 192F, 192M Modem + Transceiver								
400-420	150-174	12.5	10K8F2D	9600	?		-101 dBm (@ 12 dB SINAD ?)	deviation +/- 3KHz; packet size adjustable 1-2000 bytes; FEC (undefined type); CRC-32 error detection; packet overhead approx.125 bytes; 15 msec xmit warm up.
400-420	150-174	25	17K6F1D	19200	?		-101 dBm (@ 12 dB SINAD ?)	deviation +/- 4KHz; packet size adjustable 1-2000 bytes; FEC (undefined type); CRC-32 error detection; packet overhead approx.125 bytes; 15 msec xmit warm up.
Satel (Finland)								
Satelline 3AS/3ASd Modem + RF								
380-470		12.5, 25	F1D	19200 (25), 9600 (12.5)	?	1x10-3 @ -110 dBm	?	LCD panel, flash memory, up to 1 W; unclear whether FCC certified; specs also list BER at 1x10-3@-116 dBm
Satelline 2ASxE Modem + RF								
380-470		12.5, 20, 25	F1D	1200-9600 (25), 1200-4800 (12.5)	FSK	1x10-3 @ -115 dBm	?	packetized protocol, 135 bytes; up to 1 W; unclear whether FCC certified.

Appendix A: Integrated Radio Modems

UHF (MHz)	VHF (MHz)	Channel BW (KHz)	FCC Emission Designators	Over the Air bps	Modulation	BER	Sensitivity (12 dB SINAD)	Notes
RF Neulink								
Neulink 9600 Modem + Transceiver								
403-512		25, 12.5	16K0F1D, 16K0F2D	9600	GMSK	1x10-3 @ -110 dBm (0.7 uV)	?	configurable packet size; packetized; 16 bit CCIT CRC; unit has 16 bit ID; 2W transmitter; BER is also listed as 1x10-4 @ 0.7 uV; unclear if 12.5 kHz channels are 9600 bps
	136-174	25, 12.5	16K0F1D, 16K0F2D	9600	?	1x10-4 @ -113 dBm (0.5 uV)	?	configurable packet size; packetized; 16 bit CCIT CRC; unit has 16 bit ID; 2W or 4W transmitter; unclear if 12.5 kHz channels are 9600 bps
Ultracom Oy (Finland)								
Ultracom M02 Modem + Transceiver								
442-470		12.5, 20, 25		7200 (25k), 3600 (12.5)	Baseband FSK	1x10-3 @ -110 dBm	-113 dBm	Only certified to ETS 300 -113, not FCC certified; synthesized with fast turn on <10 msec; 5 W digitally adjustable; No processing - direct CMOS input to modulator; bps is max.
DW4485 Modem + Transceiver								
400-470		12.5		4800	Baseband FSK	1x10-2 @ -110 dBm	-113 dBm	Only certified to ETS 300 -113, not FCC certified; 2 channel crystal radio; 400 mW; No processing - direct CMOS input to modulator; +/- 1.5 kHz deviation.
Teledesign Systems, Inc								
TS4000 Radio Modem (list price UHF 2W $980, qty 1)								
380-512 (928-960)	132-174	5, 6.25, 7.5, 10, 12.5, 15, 25, 30		2000-19200	GMSK, CPFSK, 4LevelFSK	1x10-6 @ -103 dBm (typical)	?	programmable Carrier Detect -110 dBm to -60 dBm; programmable power out; up to 5 W; Hamming FEC code (12,8) (50% overhead), 16 bit interleaving, Data Scrambler; synthesized; flash programmable; transparent of AIRNET Protocol (version of MX.25?); specs also state 1x10-4 @ -103 dBm
TS2000 2400 Modem ONLY (list price $399, qty 1)								
N/A	N/A	N/A		1200 or 2400	CPMSK	1x10-5 @ 14 dB SINAD, uncoded	N/A	Modem only; full or half duplex; programmable xmit attack time; Hamming FEC code (12,8) (50% overhead), 16 bit interleaving, Data Scrambler; flash programmable; AIR NET Protocol

Appendix B: New ALERT Data Link Layer
Header Packet Only - ALERT Data

	bytes	bits	
Carrier Only	0	0	Open Squelch
Bit Sync	12	96	pattern "1100110011001100" for GMSK
Frame Sync	6	48	for strong correlation: multiple concatenated NASA sync code
Header Packet			
source address	2	16	
protocol identifier	1	8	PID = 0 for ALERT Header only transmission
reserved	1	8	
data	2	16	Sensor ID
data	2	16	Sensor Value
outer FEC Code	4	32	Reed-Solomon - shortened code error correction & detection
Convolutional Code OH	12	96	
Convolutional Code Tail	1.75	14	
Data Block			
data	0	0	(none)
outer FEC Code	0	0	Reed-Solomon - shortened code error correction & detection
Convolutional Code OH	0	0	
Convolutional Code Tail	0	0	

Over the Air Rate bit rate	3600	symbols/second, binary encoding
Convolutional Code Rate	0.50	rate=1/2
Code Constraint	7	k=7
Data Block - data bytes	0	bytes
Total Packet Size	43.75	bytes
Transmission Time	97	msecs (bit sync, frame sync, header with data)
effective Data bps	658	bps

Comparison with Current ALERT

typical 300bps ALERT	313	msecs	180	msecs preamble	133	msecs data
13+11 bits	3	bytes				
effective Data bps	73	bps				
Channel Capacity Increase	322%	in stations				
Data Rate Increase	897%	in bps				

RS code bytes required for data block size lookup table

	Data Block size	RS bytes
0	for 0-128	4
0	for 129-251	4
0	for 252-502	8
0	for > 503	16

Appendix B: New ALERT Data Link Layer
Header Packet with 28 byte Data Block

	bytes	bits	
Carrier Only	0	0	Open Squelch
Bit Sync	12	96	pattern "1100110011001100" for GMSK
Frame Sync	6	48	for strong correlation: multiple concatenated NASA sync code
Header Packet			
source address	2	16	
protocol identifier	1	8	PID = 1 for Standard Data Block
reserved	1	8	
data	1	8	data block length: # of 4 byte increments max= 1020 bytes
data	1	8	data encoding
data	2	16	destination address
outer FEC Code	4	32	Reed-Solomon - shortened code error correction & detection
Convolutional Code OH	12	96	
Convolutional Code Tail	0	0	
Data Block			
data block PID	1	8	
data	27	216	bytes are interleaved for burst error correction.
outer FEC Code	4	32	Reed-Solomon - shortened code error correction & detection
Convolutional Code OH	5	40	
Convolutional Code Tail	1.75	14	

Over the Air Rate bit rate	3600	symbols/second, binary encoding
Convolutional Code Rate	0.50	rate=1/2
Code Constraint	7	k=7
Data Block - data bytes	28	bytes
Total Packet Size	80.75	bytes
Transmission Time	179	msecs (bit sync, frame sync, header & data block)
effective Data bps	1516	bps

Comparison with Current ALERT

typical 300bps ALERT	313	msecs	180	msecs preamble	133	msecs data
13+11 bits	3	bytes				
effective Data bps	73	bps				

Channel Capacity Increase	175%	in stations
Data Rate Increase	2065%	in bps

RS code bytes required for data block size lookup table

	Data Block size	RS bytes
4	for 0-128	4
0	for 129-251	4
0	for 252-502	8
0	for > 503	16

Appendix C: Issues to be Addressed in Prototype Development

1) For the decoding, can a high performance, general purpose, low power, RISC architecture microcontroller (e.g. TI MSP430 or Mot HC12) execute the digital filtering, FSK demodulation, bit clock digital PLL and sync correlation, Viterbi (or Fano) decoding, RS decoding all simultaneously in real time? Probably not. Alternative architectures to be considered:

    - Multiple microcontrollers allowing significant power control by turning off unneeded processors.
    - A higher performance microcontroller (Atmel megaAVR microcontroller at 20-25 MIPs).
    - Low cost, low power DSP with 30-50 MIPs such as the TI TMS320x54xx series (built in “Viterbi accelerators”) and an additional low-power microcontroller.

    Example execution time for Analog Device’s older 2191 DSP running a GSM (r = $^1/_2$, k = 5) Viterbi decoder[67] is listed as 192 µsec[68] for a processing rate of 5,000 encoded bps. This is probably not sufficient performance to decode 3600 bps with k=7 or 9. On the other hand, Viterbi decoding rates for a newer communications optimized DSP, such as the TI 54x series, can easily decode the 3600 bps encoded bit rate for the new protocol in real time using only a small percentage of their computational performance. The following table shows the performance of the C54x series in terms of required MIPs for various Viterbi decoding algorithms (taken from [Hendr02]):

Standard	Data Type	Coding Rate (R)	Puncture RatE (PR)	Constraint Length (K)	Frame Size (FS)	Frame Rate (FR)	Benchmark (MIPS)
GSM	Voice	1/2	–	5	189 bits	50 Hz	0.58
	Data – 9.6	1/2	57/61	5	244 bits	50 Hz	0.75
	Data – 4.8	1/3	–	5	152 bits	50 Hz	0.53
IS–136	Voice	1/2	–	6	89 bits	50 Hz	0.46
	FACCH	1/4	–	6	65 bits	50 Hz	0.42
WLL†	Voice	1/2	2/3	7	130 bits	50 Hz	1.20
	FAX	1/2	2/3	6	190 bits	50 Hz	0.97
IS–95	Forward Voice	1/2	–	9	192 bits	50 Hz	6.49
	Reverse Voice	1/3	–	9	192 bits	50 Hz	6.57

† Wireless local loop – proprietary standard

Table A-1: Viterbi decoding computational requirements

[67] Source Code available at: http://www.analog.com/Analog_Root/static/technology/dsp/EZAnswers/codeExamples/219x_single_core.html
[68] 30,751 cycles at 160 MHz required to decode the GSM rate ½, k=4 code, 189 Class I and Class II encoded voice block, using a 4-bit quantized soft decision algorithm.

The cost for a TMS320C5402, 100 MIPs device is budgetary $5.80 (1,000 qty.), but a DSP solution based on a low end device like this is not "single chip" due to requirements for off chip EPROM and "glue" logic, unlike current embedded or RISC embedded processors. TI's recently introduced flash DSPs[69] (the highly integrated F2812, with 128K bytes of flash memory and many peripherals) is targeted for control applications, not communications, and is expensive for this application at a budgetary cost of $50.94 (1,000 qty.), making it an unlikely candidate. Other TI and Motorola flash DSPs[70] with 8K – 60K bytes and 40+ MIPS with budgetary pricing at $10 - $20, may be worthy of further evaluation, although they also are targeted at motor control applications, not communications.

2) Many third parties are selling Intellectual Property (IP) targeted at DSP communications algorithms. For reference, Imagine Technology sells an optimized convolutional code encoder and Viterbi decoder software function for the TI TMS320C5x DSP platform for $7,500.[71] Other IP is available also, such as 4i2i's Reed Solomon encoder/decoder for the 8051 microcontroller architecture and a Reed Solomon decoder for Altera Inc.'s FPGA for $3,995 and an encoder for $1,995[72]. Unfortunately, although this purchased IP is usually "royalty free," it is not "open source," and therefore the resulting ALERT protocol implementation would not be available for manufacture by any other vendors. (Although the specifications would be open.) Therefore the purchase of IP doesn't appear to be viable.

3) Most ASSP demodulators and channel decoders uncovered in research are highly integrated and targeted to a specific protocol, for instance[73]:

- STMicroelectronics' STA002 "Starman" World Space satellite radio IC with a complete QPSK demodulator, Nyquist root raised cosine filter, Viterbi $r = {}^1/_2$, $k = 7$ soft decision decoder, de-interleaver, and Reed-Solomon (255,223) decoder.

- Atmel's TSS902E, a high performance (10Mbps) RS decoder, de-interleaver and Viterbi convolutional decoder for Digital Video Broadcast Standard.

- CML's CMX909B, a Reed-Solomon encoder and decoder, with interleaver and de-interleaver and modulator and demodulator for Mobitex applications.

- CML's CMX 919/929 "data pump" ICs with a 4 level RC FSK modem, with a convolutional coder and Viterbi encoder for RD-LAP applications. This may, in fact be a good option if the RD-LAP protocol is adequate to meet the ALERT requirements, although unfortunately the MDC-4800 and RD-LAP protocols (including the air interface) are proprietary to Motorola, although licensable.

- A general purpose Viterbi coder/decoder (albeit targeted at very high throughput) such as the Intel STEL 2030C may be suitable, but it is an old IC and not readily available.

---

[69] The TMS320F2812 at 150 MIPs. See http://focus.ti.com/docs/prod/productfolder.jhtml?genericPartNumber=TMS320F2812&pfsection=features

[70] See the TI TMS320F241/243 and TMS320LF2406A and the Motorola DSP56F807 and DSP56F8xx series in general.

[71] Email price quote from Scott Kroeker, Director of Business Development, Imagine Technology, Lincoln, NE, Tel. 402-472-3321 ext. 4, email: skroeker@imaginetechnology.net web page: www.imaginetechnology.net

[72] See http://www.altera.com/corporate/news_room/releases/releases_archive/2000/pr-ninecores.html as of October 2002.

[73] See also: PraireComm's PCI3700 TIA/EIA-136/GSM/GPRS Integrated Baseband Processor, with GFSK or Pi/4 DQPSK modulator/demodulators, Viterbi hardware decoder, GSM Fire/Cipher Accelerator, with additional A/D's D/A's, UARTs, Codecs, and an ARM 32 bit RISC processor, (they may as well have included the kitchen sink), all extremely low power.

The current state of the art for non-ASSP modulation and channel coding is a DSP, FPGA (Xilink or Altera) or embedded high performance CPU (Power PC, Intel, ARM) based.

4) A general purpose GMSK modem chip is available from CML, the CMX589A. Although it is limited to a max BT of 0.5 it is an inexpensive part with demonstrated good BER performance and a low cost ($12). It is a viable candidate. It is probably interoperable with the G3RUH Amateur standard.

References

[Ander95] Jørgen Bach Andersen, Theodore S. Rappaport and Susumu Yoshida, Propagation Measurements and Models for Wireless Communications Channels, IEEE Communications Magazine, vol. 33, no. 1, pp. 42 – 49, January 1995

[Ander91] John B. Anderson and Carl-Erik W. Sundberg, “Advances in Constant Envelope Coded Modulation,” *IEEE Communications Magazine*, vol. 29, no. 12, pp. 36 – 45, December 1991.

[Aulin81a] Tor Aulin and Carl-Erik Sundberg, “Continuous Phase Modulation – Part I: Full Response Signaling,” *IEEE Transaction on Communications*, vol. 29, no. 3, pp.196-209, March 1981.

[Aulin81b] Tor Aulin and Carl-Erik Sundberg, “Continuous Phase Modulation – Part II: Partial Response Signaling,” *IEEE Transaction on Communications*, vol. 29, no. 3, pp.210-225, March 1981.

[Berle87] Elwyn R. Berlekamp, Robert E. Peile, Stephen P. Pope, “The Application of Error Control to Communications,” IEEE Communication Magazine, vol. 25, no. 4, pp. 44-57, April 1987.

[Berro96] Claude Berrou and Alain Glavieux, “Near Optimum Error Correcting Coding and Decoding: Turbo-Codes,” *IEEE Transactions on Communications*, vol. 44, no. 10, pp. 1261-1271, October 1996.

[CCSDS01] “Recommendation for Space Data System Standards, Telemetry Channel Coding, Blue Book, CCSDS 101.0-B-5,” June 2001, Consultative Committee for Space Data Systems.

[Chaya93] Naftali Chayat and Michael Rothenberg, “Simulation results for several WLAN Modulation Methods,” IEEE 802.11 Working Group Submission, 1993 (from Lannair Inc., Tel-Aviv, Israel), available at: ftp://ftp.cs.utwente.nl/pub/doc/802.11/ibmpc/1993_docs/1193135.doc as of September 2002.

[Chuan87] Justin C-I Chuang, “The Effect of Time Delay Spread on Portable Radio Communications Channels with Digital Modulation,” IEEE Journal on Selected Areas in Communications, col. SAC-5 no. 5, pp 879-889, June 1987.

[Dolin88] S. Dolinar, “A New Code for Galileo,” *TDA Progress Report 42-93*, January-March 1988, Communications Systems Research.

[Dries00] Peter F. Driessen, “Prediction of Multipath Delay Profiles in Mountainous Terrain,” IEEE Journal on Selected Areas in Communications, vol. 18, no.3, pp. 336-346, March 2000.

[Euro94] “Navigation Data Integrity Model, ECC Report 276,” EEC Task AT57, EATCHIP – NAV.ET1.ST15, European Organization fro the Safety of Air Navigation, 1994.

[Finton] Mike Finton, Ph.D., “Principles of Digital Modulation,” a presentation from the Telecommunications Research Lab, Toshiba Research Europe Ltd.

[Flemi01] Chip Fleming, “A Tutorial on Convolutional Coding with Viterbi Decoding, Example Simulation Results,” October 8, 2002, by Spectrum Applications Corp. available October 2002 at: http://pw1.netcom.com/~chip.f/viterbi/tutorial.html

[Fleur96] Bernard H. Fleury and Peter E. Leuthold, Radiowave Propagation in Mobile Communications: An Overview of European Research, IEEE Communications Magazine, vol. 34, no. 2, pp. 70 – 81, February 1996

[Guma98] Charles Gumas, “Turbo codes rev up error-correcting Performance,” *Personal Engineering & Instrumentation News*, January 1998, pp. 61-66

[Hendr02] Henry Hendrix, “Viterbi Decoding Techniques for the TMS320C54x DSP Generation,” *TI Application Report SPRA071A*, Texas Instruments, January 2002

[Intel02] Intel Corporation, Cable Network Operations, “STEL-2030C IC Data Sheet”, January 7, 2002, available October 2002 at: ftp://download.intel.com/design/digital/datashts/24540001.pdf

[Ishiz80] Mitsuru Ishizuka and Kenkichi Hirade, “Optimum Gaussian Filter and Deviated-Frequency-Locking Scheme for Coherent Detection of MSK,” *IEEE Transactions on Communications*, vol. 28, no. 6, pp. 850 – 857, June 1980.

[Iwan95] Yasunori Iwanami, “Performance of Sequence Estimation Scheme of Narrowband Digital FM Signals with Limiter-Discriminator Detection,” *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 2, pp. 310 – 315, February 1995.

[Karn95] Phil Karn, “Convolutional Decoders for Amateur Packet Radio,” presented at the 1995 ARRL Digital Communications Conference. See http://www.tapr.org/tapr/html/dccconf.html for ARRL/TARPA information; see http://www.ka9q.net/papers/cnc_coding.html for actual paper. (As of 10/2002)

[Karn02] Phil Karn, source code for convolutional en/decoders and Reed-Solomon en/decoders are available at http://www.ka9q.net/code/fec/ as of October 2002.

[Korn91] Israel Korn, GMSK with Limiter Discriminator Detection in Satellite Mobile Channel, IEEE Transactions on Communications, vol. 39, no. 1, pp. 94 – 101, January 1991

[Lee99] C. S. Lee, T. Keller and L Hanzo, “Turbo-Coded Digital Video Broadcasting for Mobile Environments,” paper presented at EURASIP 1999 Conference, Krakow, June 1999, available: http://www-obile.ecs.soton.ac.uk/cs/papers/DVB-T_turbo_paper.pdf as of October 2002.

[Lin83] Shu Lin, Daniel J. Costello Jr., *Error Control Coding: Fundamentals and Applications*, Prentice-Hall, Inc., Engelwood Cliffs, New Jersey 1983

[Ma80] Howard H. Ma, “The Multiple Stack Algorithm Implemented on a Zilog Z-80 Microcomputer,” *IEEE Transactions on Communications*, vol. 28, no. 11, pp. 1876 – 1882, November 1980.

[Miller88] James R Miller, “9600 baud Packet Radio Modem Design,” *Proc. 7th ARRL Computer Networking Conference*, pp. 135-140, October 1988

[Miller91] James R Miller, “The Shape of Bits To Come,” Amsat-UK’s Oscar News, No. 88, pp. 29-37, April 1991. Also available at: www.amsat.org/amsat/articles/g3ruh/108.txt as of May 2002.

[Nesko00] Aleksandar Neskovic, Natasa Neskovic, and George Paunovic, “Modern Approaches in Modeling of Mobile Radio Systems Propagation Environment,” *IEEE Communication Surveys*, Third Quarter 2000, pp. 2-12.

[Pasup79] Subbarayan Pasupathy, “Minimum Shift Keying: A Spectrally Efficient Modulation,” *IEEE Communications Magazine,* July 1979.

[Petro00] Rade Petrovic and Andtrea F. Molisch, "Multipath Effect of FSK with Frequency-Discriminator Detection," IEEE Transactions on Vehicular Technology, vol. 49, no. 3, pp. 856-862, May 2000.

[Proak01] John G. Proakis, Digital Communications, Fourth Edition, McGraw-Hill Higher Education, McGraw-Hill Companies, New York, NY, 2001

[Rappa96] Theodore S. Rappaport, *Wireless Communications, Principles and Practice*, Prentice Hall PTR, A Simon & Schuster Company, Upper Saddle River, New Jersey, 1996

[Rohd01] Ulrich L. Rohde and Jerry C. S. Whitaker, *Communications Receivers: DSP, Software Radios, and Design, 3rd Edition*, McGraw-Hill Companies, New York, NY, 2001.

[Schul98] Christian Schuler, "Christian Schuler's Forward Error Correction (FEC) Page," Reed-Solomon 'C' source, copyrighted 1998, but available for "open source" use, developed at GMD-FOKUS (the Fraunhofer Institute for Open Communication Systems) and available at: http://home.arcor.de/christianschuler/fec_sw.html as of October 2002.

[Shan48] C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, vol. 27, pp. 379-423, July 1948, and vol. 27, pp.623-656, October 1948.

[Silbe93] Nathan Silberman, "Proposal for a modulation technique for Frequency Hopping spread Spectrum PHY Standard, Wireless Access Methods and Physical Layer Specifications," by California Microwave, Sunnyvale, CA, May 1993, IEEE 802.11 doc:IEEE P802.11-93/84, available at:
http://grouper.ieee.org/groups/802/11/Documents/DocumentArchives/1993_docs/1193084.DOC as of October 2002.

[Sklar01] Bernard Sklar, *Digital Communications, Fundamentals and Applications*, Second Edition. New Jersey: Prentice Hall P T R, 2001.

[Sund86] Carl-Erik Sundberg, "Continuous Phase Modulation," *IEEE Communications Magazine*, Vol.24, No.4, April 1986.

[Tjhu70] Tjeng T. Tjhung and Paul H. Wittke, "Carrier Transmission of Binary Data in a Restricted Band," IEEE Transactions on Communications Technology, vol. COM-18, no. 4, pp 295 -304, August 1970.

[Valen01] Reinaldo A. Valenzuela, Director, Wireless Research, Bell Labs (Lucent), "Broadband Wireless Tutorial," a course presentation dated 3/20/2001.

[Wang02] Charles Wang, Dean Sklar, and Diana Johnson, "Forward Error-Correction Coding," *CrossLink, The Aerospace Corporation Magazine,* Volume 3, Number 1 (Winter 2001/2002, available at http://www.aero.org/publications/crosslink/winter2002/04.html as of October 2002.

[Weck88] J. P. De Weck, P. Merki, and R. W. Lorenz, "Power delay profiles measured in mountainous terrain," in Proc. IEEE Veh. Tech. Conf., pp. 105–112, 1988