Asymptotic Throughput of the FED-STD-1052 Data Link Protocol

Eric E. Johnson

New Mexico State University
Las Cruces, NM

ABSTRACT

The past decade has seen rapid evolution of data communication technologies specially adapted to the challenges of HF skywave channels. The 1980s saw the development of advanced HF modems that achieve significant processing gains through adaptive equalization and advanced coding. The recently standardized HF data link link protocol of FED-STD-1052 and MIL-STD-187-721C effectively uses the capabilities of these new modems to achieve reliable data delivery at rates well above those previously attainable.

This paper presents a straightforward mathematical model for predicting the throughput of user data that can be achieved by the HF Data Link Protocol (HFDLP) and the single-tone HF data modem over a wide range of channel conditions. This model is based upon measurements of modem performance, and is validated by measurements of protocol throughput. The model can be implemented in a spreadsheet so that HF system designers can easily predict and optimize the performance of their systems.

Recent measurements of other popular HF protocols are included for comparison.

INTRODUCTION

HF skywave channels are among the most challenging in communications system design, due to large fluctuations in the channel impulse response on every conceivable time scale:

- Noise impulses with nanosecond to microsecond durations.
- Multipath effects on the scale of milliseconds.
- Fading that lasts seconds to minutes.
- Diurnal and seasonal variations in usable channels from hour to hour and day to day.
- Solar cycle effects that vary from year to year.

The 1980s saw the development of advanced HF modems that achieve significant processing gains through adaptive equalization and advanced coding. These techniques can overcome many of the short-term effects of the channel (milliseconds to seconds). Automatic link establishment (ALE) [1] and sounding can be used by automated HF controllers to adapt to longer-term changes in propagation (minutes or longer). However, none of these techniques is efficient in coping with fades that last more than a few seconds. This time scale is the realm of an ARQ data link protocol. The recently standardized HF Data Link Protocol (HFDLP) of FED-STD-1052 [2] and MIL-STD-187-721C [3] effectively uses the capabilities of advanced modems and ALE to achieve reliable data delivery at rates well above those previously attainable.

This paper presents a straightforward mathematical model for predicting the throughput of user data that can be achieved by the HFDLP and the single-tone HF data modem over a wide range of channel conditions. This model is based upon measurements of modem performance, and is validated by measurements of protocol throughput. The model can be implemented in a spreadsheet so that HF system designers can easily predict and optimize the performance of their systems.

The paper begins with an overview of the standard single-tone modem and of the HFDLP.

BACKGROUND

Single-Tone HF Data Modem

A wide range of modulations has been used to convey information over the HF channel. The ALE modem [1] is an 8-ary frequency-shift keyed (FSK) design, and uses a long 8 ms baud to overcome multipath of up to 4 ms. However, this slow FSK limits the data rate of the ALE modem to 375 bps. The "long baud" philosophy is also evident in the "parallel-tone" modems, which employ slow phase-shift keying (PSK) of a large number of tones (for example, 39 tones [4] or 52 tones [5]) within the audio passband to achieve data rates of up to 7200 bps.

A rather different approach is used in the "single-tone" HF data modem: with but a single PSK carrier tone, high data rates require fast modulation. To overcome the inevitable intersymbol interference introduced by the multipath channel, these modems employ adaptive channel equalizers and feedback-guided decoding, along with substantial forward error correction (FEC) to clean up whatever decoding errors slip through. Interleaving is used to reduce the burstiness of errors introduced by fading, so that the FEC decoders are less likely to be overwhelmed.

Such a modem is described in MIL-STD-188-110A and FED-STD-1052. Standard interleaver settings are 4.8 s, 0.6 s, and 0 s (no interleaving), with coded data rates ranging from 75 bps to 2400 bps (4800 bps with no FEC). The training sequences for this modem contain codes indicating the data rate and interleaver settings in use. The resulting ability of a demodulator to automatically track changes in these parameters frees the modulator to dynamically vary the data rate and interleaver to compensate for changes in channel conditions.

Table 1 lists the performance required by the standards for implementations of this modem design. Channels with one path are non-fading channels with Gaussian noise, and typify surface channels. The channels with two paths are intended to represent skywave channels. ITU-R [6] defines a channel with two paths, 0.5 ms multipath delay, and 0.1 Hz fading bandwidth (or Doppler spread) as a “good channel,” and a two-path channel with 2 ms multipath delay and 1 Hz fading bandwidth as a “poor channel.” It is interesting to note that the PSK HF data modem is required to operate well in channels that are substantially worse than the ITU-R “poor channel.”

A state-of-the-art implementation of this design achieved the BER performance in Table 2 over a “very poor” fast fading channel with 2 ms multipath delay and 2 Hz Doppler spread, using a 4.8 s interleaver.

HFDLP

The HFDLP [2, 7] is a selective repeat ARQ protocol with the ability to adaptively vary several parameters in response to changing channel conditions. A transmission usually consists of a data series, containing several data frames, or a single control frame. Every frame contains a CRC.

Before data transfer commences, HFDLP terminals exchange control frames to negotiate the number of data bytes per data frame (56 to 1023), the number of data frames per data series (1 to 255), and a few other characteristics of the data transfer procedure. In its simplest case, the data transfer then proceeds as follows:

- The “transmit terminal” sends a data series, using the negotiated frame and series sizes.
- The “receive terminal” responds with a control frame which contains individual acknowledgement bits for each frame in the received series. Received frames are acknowledged only if their CRCs are correct.
- The transmit terminal prepares and sends a new data series, which begins with any frames from the preceding series not acknowledged by the receive terminal.
- This cycle is repeated until an entire message is successfully transferred or the link fails.

Table 1: Required Modem Performance

User Data bps	Channel Paths	Multipath (ms)	Fading BW (Hz)	3 kHz SNR (dB)	Coded BER (maximum)
4800	1	—	—	17	1e-3
4800	2	2	0.5	27	1e-3
2400	1	—	—	10	1e-5
2400	2	2	1	18	1e-5
2400	2	2	5	30	1e-3
2400	2	5	1	30	1e-5
1200	2	2	1	11	1e-5
600	2	2	1	7	1e-5
300	2	5	5	7	1e-5
150	2	5	5	5	1e-5
75	2	5	5	2	1e-5

In addition to this simple case, the HFDLP contains provisions for link establishment, forward and reverse preemption of messages in progress, message resumption after preemption or link loss, adaptation of frame size, series size, data rate, and interleaver setting, and several other functions. For further details, see [2, 7].

Communication system architects seeking to design this protocol into HF networks will certainly need to know what data throughputs may be achieved under a range of channel conditions. In many cases, measurements for the specific combinations of channel condi-

Table 2: Measured Modem Performance

3 kHz SNR (dB)	User Data bps	Coded BER (measured)
-6	75	2e-2
-5	75	2e-3
-4	75	3e-4
-3	75	5e-5
-2	150	2e-2
-1	150	2e-3
0	150	1e-4
1	300	3e-3
2	300	3e-4
3	300	3e-5
4	600	6e-4
5	600	6e-5
6	1200	1e-2
7	1200	3e-3
8	1200	6e-4
9	1200	7e-5
10	2400	2e-2
11	2400	7e-3
12	2400	2e-3
13	2400	5e-4

tions and protocol parameters of interest will not be available. Development of simulation experiments may not be practical under the time constraints faced by many designers. An analytical model of this protocol will be quite useful in such situations.

HFDLP THROUGHPUT ANALYSIS

This section presents a simple mathematical model of the HFDLP and the PSK HF data modem operating over HF skywave channels. Although a number of simplifying assumptions are made in this model, it nevertheless produces satisfactory predictions of the achievable throughput of this protocol/modem combination.

The model uses a number of parameters, which are defined in groups below.

Protocol Characteristics (constants)

b_h	frame overhead bits per data frame (excluding sync bits): 72
b_c	bits per control frame (excluding sync bits): 520
b_s	sync bits per frame (excluding transmission overhead bits): 16
b_x	overhead bits per transmission (first 8 sync bits, modem flush): 184

Operating Parameters

n_d	data frames per data series
b_u	user data bits per data frame
b_d	total bits per data frame $= b_u + b_h$
r_d	data rate for data series
r_c	data rate for control frames
v_d	interleaver for data series
v_c	interleaver for control frames
t_p	processing time in HFDLP controller between frame arrival and response
t_{ld}	modem latency for data series $= v_d$
t_{lc}	modem latency for control frame $= v_d$

Metrics

p_e	probability of a bit error after FEC; p_e = BER.
p_d	probability that a data frame is received error-free
p_c	probability control frame is received error-free
t_{ts}	transmit (on-air) time for data series
t_{tc}	transmit (on-air) time for control frame
n_c	expected number of times each control frame must be sent for error-free reception
t_{ds}	total elapsed time for data series $= t_p + t_{ld} + t_{ts}$
t_{ack}	total elapsed time to complete ACK phase
t_{cyc}	total elapsed time for data series and ACK phase $= t_{ds} + t_{ack}$
X_u	user data throughput (bps)

Model Derivation

For a very long file transfer, the total time to convey the file over the HF channel is dominated by the alternating data series and acknowledgment transmissions described above. That is, the overhead involved in starting the file transfer is relatively insignificant. The model derived here therefore ignores this startup overhead, and therefore should be considered to predict the asymptotic throughput approached by long file transfers. For shorter messages, it is a simple matter to add an additional overhead term to the model, as discussed later.

Other simplifying assumptions used here are that errors in the bit stream received from the modem are independently distributed, and that control frames containing errors prompt an immediate response from the terminal that detects the corrupted control frame.

Although the errors produced by HF skywave channels are quite bursty, the errors after de-interleaving and FEC decoding in the modem will be relatively rare and more uniformly distributed than the channel errors. Note that the assumption of independent errors is pessimistic, because increased burstiness in the error stream will tend to localize errors in fewer frames, leading to higher throughput than for independent errors.

The assumption of immediate responses to corrupted control frames, however, is optimistic. When a control frame is lost or ignored, rather than detected as corrupt, an HFDLP terminal will wait for expiration of a timeout before retransmitting. This would lengthen the time to complete the ACK phase of the protocol, and thereby decrease throughput.

The model permits the data rates and interleaver lengths for data series and control frames to be varied independently. This is assumed to occur without negotiation phases during data transfer.

Throughput Model

The operation of the protocol is modeled as an integral number of the data-series/ACK cycles described earlier. Between each reception and subsequent transmission the protocol requires a minimum turn-around time (t_p) of 1 second. Another delay within each station is the latency within the modem between the arrival of the last bit of a received transmission and the delivery of the last data bit to the HFDLP processor. This latency (t_{ld} or t_{lc}) occurs as the received data block is de-interleaved. (The latency is exactly equal to the interleaver length when the data path from the modem to the processor runs at the on-air data rate, which is assumed here.)

The modem must always send an integral number of interleaver blocks, plus an initial training sequence for each transmission that is the same length as an interleaver block. Thus, the time to generate and send each data series is $t_{ds} = t_p + t_{ld} + t_{ts}$, where

$$t_{ts} = v_d \left(1 + \left\lceil \frac{n_d \left(b_d + b_s\right) + b_x}{r_d \, v_d} \right\rceil \right)$$

is the time that the transmitter is on the air sending the training sequence followed by the data series.

Similarly, the time to send a control frame (and its preceding training sequence) is

$$t_{tc} = v_c \left(1 + \left\lceil \frac{b_c + b_s + b_x}{r_c \, v_c} \right\rceil \right)$$

Of the n_d frames in each data series, some may arrive with errors. With the assumption of independence in the error locations, the probability that any data frame is error-free is simply

$$p_d = \left(1 - p_e\right)^{b_d}$$

Similarly, for an error-free control frame we have

$$p_c = \left(1 - p_e\right)^{b_c}$$

and the expected number of times that a control frame must be sent is

$$n_c = p_c^{-1}$$

Successful completion of an ACK cycle occurs when an error-free data ACK control frame is received by the transmit terminal. This may be preceded by a series of corrupted data ACK receptions, which prompt data-ACK-request control frame transmissions from the transmit terminal. The receive terminal is assumed to respond immediately to these frames, whether or not they are received error-free. Except for preemption or renegotiation requests (neither of which is considered here), the only cause for the arrival of a control frame from the transmit terminal is the loss of a data ACK, and the receive terminal responds accordingly. Since the time to prepare and send a control frame is $t_p + t_{lc} + t_{tc}$, the time to complete the ACK phase is $t_{ack} = (t_p + t_{lc} + t_{tc}) \bullet (1 + 2\,(n_c - 1))$.

A complete cycle then consumes $t_{cyc} = t_{ds} + t_{ack}$. Each data series conveys $(p_d \, b_u \, n_d)$ error-free bits, so the user data throughput is

$$X_u = \frac{p_d \, b_u \, n_d}{t_{cyc}} \text{ bps.}$$

Example Application

The model described above can be easily implemented in a spreadsheet, as shown in Table 3. For a simple implementation, the cells in italic type are user entries, with the other cells calculated by the model. A more powerful implementation employs a table of BER values as a function of SNR, data rate, and interleaver, and automatically adjusts the data rates and interleavers to optimize throughput, given only the SNR.

Poor Channel Throughputs

Using the measurements of modem performance in Table 2 and the model described above to predict the performance the HFDLP/PSK data modem combination yields the results in Table 4. In each case, the interleaver for data series was set to 4.8 s, while the interleaver for control frames was set to 0.6 s. As a consequence, the error rate for control frames was approximately an order of magnitude greater than for data series at the same SNR and data rate.

For each SNR value, the data rates, data frame sizes, and data series lengths were selected for peak throughput, with a constraint that transmissions were not allowed to exceed ten minutes.

Most results in Table 4 show an approximate doubling in throughput with each 3 dB increase in SNR, as expected: each 3 dB improvement permits a doubling in the modem speed while maintaining an error rate on

Table 3: Example Throughput Computation

Very Poor Channel **9 dB SNR**	Data Series	Control Frames	Units	Notes
BER (after modem FEC)	*7e-5*	*7e-4*		(est for 9 dB SNR 2 ms 2 Hz)
User data	*256*		bytes	
Frame size	2120	520	bits	(excl preamble, frame sync)
Frames/Series	*255*			
Pr[frame error-free]	0.862	0.695		assuming random errors
Data rate	*1200*	*1200*	bps	
Interleaver	*4.8*	*0.6*	s	
Transmit time	460.8	1.2	s	on air
Turnaround time	1.0	1.0	s	
Good bytes/series	56,277			
E[# times sent]		1.4		
Total series time	471.9		s	incl data series, ctrl frame(s)
User data throughput	954		bps	no negotiation btwn series

the order of 10^{-5}. However, the model predicts only a small gain in user data throughput when the SNR improves from 9 dB to 12 dB. Because of the high error rate of the modem when running at 2400 bps in the 12 dB SNR channel, it must operate at 1200 bps, and the gain in throughput at 12 dB is principally due to reduced retransmissions of control and data frames. From 12 to 15 dB, the modem rate can again double, and the model consequently predicts a near doubling in user data throughput as well.

Table 4: Predicted Asymptotic Throughputs
(Very Poor Channel)

3 kHz SNR (dB)	Throughput (bps)	Optimum Data Rate (bps)
0	111	150
3	255	300
6	550	600
9	954	1200
12	1139	1200
15	2177	2400

Modification for Short Messages

When the HFDLP is used to convey short messages (e.g., command and control applications rather than file transfers), the above model will prove optimistic, because of non-negligible delays in setting up the link. When the message to be sent consumes N data transfer cycles, and a period t_{setup} is required to bring the HFDLP link into operation, the user data throughput is reduced to

$$X_u = \frac{p_d \, b_u \, n_d}{t_{cyc} + \frac{t_{setup}}{N}} \quad \text{bps.}$$

VALIDATION OF THE MODEL

Some measurements of HFDLP performance are already available. Measurements of an implementation that operates similarly to the model (e.g., no negotiation between data series) are shown in Figure 1 for comparison to the prediction of the model. The predictions from the model do not reflect adjustment for finite message size, while the measurements were collected for file sizes ranging from 1000 to 50,000 bytes. The principal source of the difference between the measured and predicted throughputs is the finite message lengths used in the measurements.

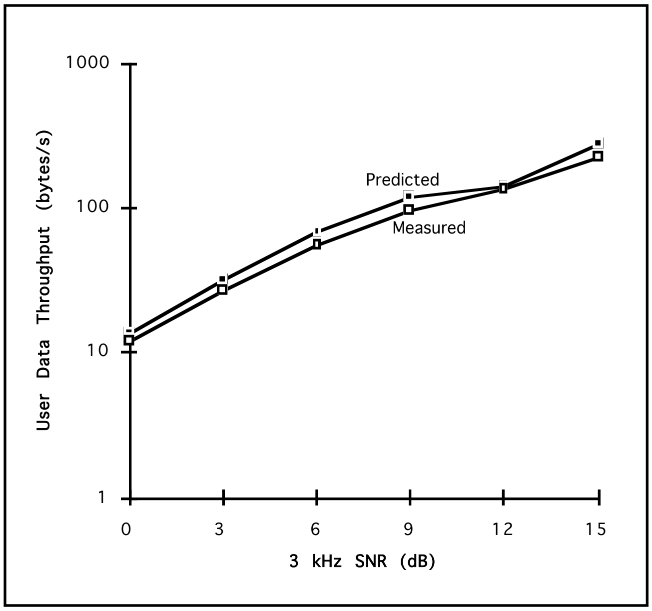

Figure 1: Comparison of Measured and Predicted Throughputs

HF PROTOCOL COMPARISON

The Institute for Telecommunication Sciences (ITS) of the Commerce Department recently measured the performance of several HF data link protocols that are popular with radio amateurs [8]. The cost of the equipment that implements these protocols is often lower than that of current implementations of the PSK HF data modem and HFDLP controllers, so it is interesting to compare the relative performance of the systems (see Figure 2).

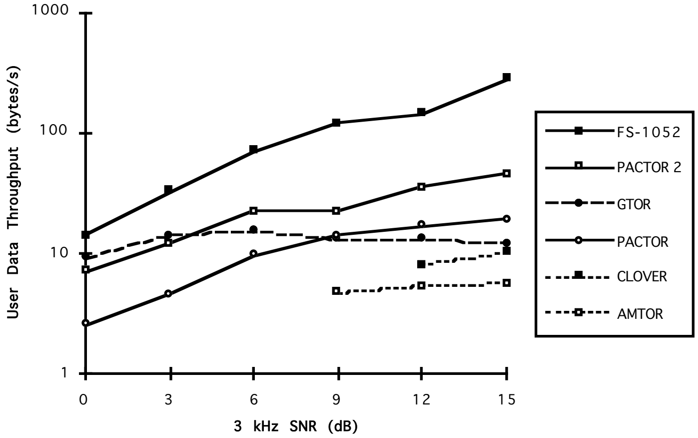

Figure 2: Throughput Comparison of HF Protocols (Poor to Very Poor Channels)

The throughputs shown are measurements of the asymptotic throughputs for the amateur radio protocols for the ITU-R poor channel (2 ms multipath and 1 Hz fading), and the predicted throughputs for the HFDLP over a somewhat worse channel (2 ms and 2 Hz). The HFDLP throughput is always at least a factor of two greater than the other protocols, and is roughly an order of magnitude faster over voice-quality channels (SNR ≥ 9 dB). The same data is presented using a linear scale for throughputs in Figure 3.

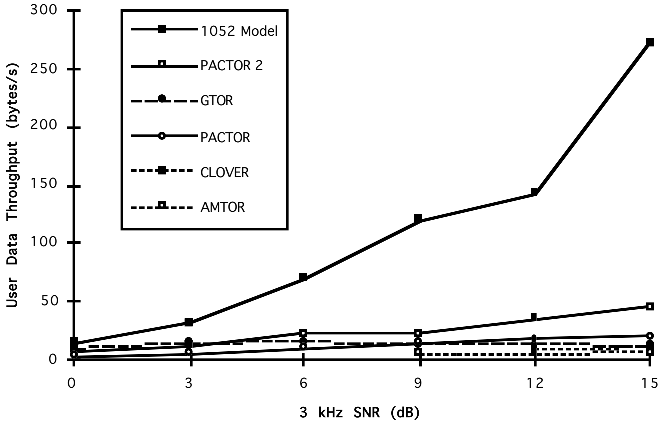

Figure 3: Linear-Scale Comparison of Throughputs

Conclusion

The simple model of the MIL/Federal standard HF Data Link Protocol presented here provides reasonably accurate predictions of the throughput that can be obtained from this protocol under various channel conditions, without resort to simulation or laboratory or field measurements. The model can be implemented in a spreadsheet for use by system engineers as they design automated HF networks using this new technology. Detailed simulation is required only for evaluation of alternative implementations of the protocol.

The performance of the HFDLP/PSK combination is seen to be substantially higher than that of the less costly systems in use by the amateur radio community.

Acknowledgments

The author is grateful to the researchers and engineers in government and industry who have provided the measurements used in this paper, and who graciously consented to their publication here.

References

1. MIL-STD-188-141A, “Interoperability and Performance Standards for Medium and High Frequency Radio Equipment,” Notice 2, 10 September 1993.
2. Proposed Federal Standard 1052, Telecommunications: HF Radio Data Modems
3. MIL-STD-187-721C, “Interface and Performance Standard for Automated Control Applique for HF Radio,” 30 November 1994.
4. MIL-STD-188-110A, “Interoperability and Performance Standards for Data Modems,” 30 September 1991.
5. Cook, S., *et al.*, “DSTO 52-Tone Modem,” Australian Defence Science and Technology Office. Private communication, 1995.
6. CCIR Recommendation 520, Use of High-Frequency Ionospheric Channel Simulators.
7. Johnson, E.E., and the HFIA, “HF Data Link Protocol,” Technical Report ASQB-94142, U.S. Army Information Systems Engineering Command, Ft. Huachuca, AZ, July 1994. Also available from `ftp://atanasoff.nmsu.edu/pub/hf/pubs/hfdlp/hfdlp.ps`.
8. Redding, C., “HF Packet Radio Modem Protocol Testing,” reported at HF Industries Association meeting, San Diego, CA, 25 January 1996. To appear in *QST*.