ARQ PROTOCOL FEATURES FOR HF RADIO LINKS

Eric E. Johnson
New Mexico State University

ABSTRACT

The explosion of interest in using Internet-style applications over wireless links prompts investigation of appropriate link-layer protocols that will most efficiently support these application and transport protocols. In this paper, we evaluate the performance impact of key features of ARQ protocols for wireless links with Internet-style workloads. The investigation is focused on high-frequency radio links, with some generalization to wireless applications throughout the spectrum.

1. INTRODUCTION

The key challenges of wireless channels that must be addressed in supporting data applications are listed below, along with closely corresponding mitigation approaches.

<u>Challenge</u>	<u>Response</u>
Low signal-to-noise ratio and Multipath fading channels	Robust modem waveforms, with forward error correction (FEC) and interleaving
Fades and interference that overwhelm FEC	Automatic repeat request (ARQ) data link protocols
Limited channel capacity	Prioritization, flow control

Some frequency bands also experience longer-term variation in channel characteristics (e.g., the diurnal and 11-year cycles that affect ionospheric channels), which requires additional adaptive techniques if data transport is to be consistently reliable.

From the perspective of the transport and application layers, these physical and link layer techniques trade latency for reliability, often introducing wide *variations* in latency as well. Such latency effects can interact strongly with the timers used by various transport and higher-layer protocols, resulting in significant degradation in throughput over wireless networks [1, 2].

In this paper, we examine features of wireless data link ARQ protocols to identify those that are most effective in supporting Internet-style applications over wireless links.

1.1 Characteristics of Internet-Style Communications

Application protocols (e.g., HTTP, FTP, and SMTP) may exchange many short commands and responses before each large file transfer (web page, file, or email message). For example, in the course of transferring a single email message, an SMTP client will send at least four short SMTP commands (typically 20 to 40 bytes each) and receive a similar number of short replies from the SMTP server. This high ratio of short to long messages is an important characteristic of Internet applications for the designer of a suitable wireless ARQ protocol.

Another common feature of internet application protocols is use of the Transmission Control Protocol (TCP) to provide a reliable stream transport service to peer entities at distant hosts. TCP breaks long application messages into segments with a fixed maximum size, and employs sliding-window flow control along with adaptive retransmission timeouts to reliably transport these segments among application entities.

The adaptive timeout mechanism of TCP is of particular interest in this study. TCP maintains an estimate of the round-trip time (RTT) on each end-to-end path, and retransmits segments when acknowledgements fail to arrive within a small multiple of its RTT estimate. When a retransmission is required, TCP responds to the possibility that the RTT has increased by increasing its RTT estimate. Of course it is also possible that the segment or its acknowledgement was lost due to network congestion, so TCP also responds to segment timeouts by reducing the rate of sending segments into the network.

As documented elsewhere [1, 2], however, the adaptive mechanisms in TCP are not optimum for wireless networks, in which link error rates are much higher than in the wireline subnetworks. Absent an appropriate wireless link ARQ protocol (or other error correction) to hide link errors from TCP, throughput can drop dramatically in response to packet loss-induced timeouts.

Both TCP and the User Datagram Protocol (UDP) call upon the Internet Protocol (IP) for routing their data through tandem connections of subnetworks, ranging from high-speed local area networks to low-bandwidth wireless networks. IP provides a "best effort" datagram service, and is generally insensitive to latency in the subnetworks.

1.2 Wireless ARQ Operation

The clients of a wireless ARQ protocol (e.g., IP) call upon the ARQ entity to deliver datagrams reliably over a link. Just as for TCP, reliability is achieved by retransmissions. However, the timeouts and other behavior of a wireless ARQ protocol can be matched to known characteristics of a specific wireless channel for more efficient operation.

A typical sequence of events in sending a datagram is a link setup handshake (if necessary to establish a data link connection), followed by cycles of alternating forward data transmissions and acknowledgements from the destination. Each such cycle on a wireless link requires that the roles of physical transmitter and receiver be reversed twice. During each "link turnaround," some delay is introduced as radio-frequency circuitry stabilizes, modems acquire and synchronize to new signals, interleavers are filled, and so on. This added cost for link turnarounds is a distinguishing characteristic of wireless links, especially for HF radio links whose turnaround times range up to tens of seconds.

Latency in wireless links also arises from the limited bandwidth of wireless channels, the time required to set up a wireless link, and the occasional delays introduced by link-layer retransmissions. Delays due to low data rate may render application protocols with fixed timeouts unusable over especially low-bandwidth channels. Long delays are readily accommodated by the adaptive timeouts in TCP, but ARQ protocols that introduce large latency *variations* may cause substantial performance loss in TCP as it attempts to adapt to such an unstable channel.

In this paper, we examine (via simulation) the effects on Internet-style communications of several features of wireless ARQ protocols. The study focuses on HF radio links as an extreme case, followed by generalization to other wireless links. The next section introduces the features examined in this study, followed by a description of the experimental design, the results, analysis of the effects of each feature, and conclusions from this study.

2. WIRELESS ARQ PROTOCOL FEATURES

The following features of ARQ protocols for wireless links were selected for study here:

- ARQ Mode: Selective Repeat *versus* Stop and Wait.
- Precise frame timing.
- Modem: code-combining (integrated) *vs.* stand-alone.
- Simplex *versus* duplex client data flow on the link.

Various combinations of these features are included in current-generation (STANAG 5066) and next-generation (3G) ARQ protocols for HF radio links:

Feature	5066	3G LDL	3G HDL
Selective repeat	•		•
Precise frame timing		•	•
Code combining		•	•
Duplex client data	•		

ARQ mode of operation. In stop-and-wait ARQ, a single data packet is sent in each forward transmission, and the sender then awaits an acknowledgement. Since the acknowledgement carries only a single bit of information, it can be encoded very robustly (to boost protocol reliability) without significantly lengthening the ARQ cycle time.

Selective repeat ARQ, on the other hand, sends several data packets per forward transmission, sharing the link turnaround burden. The return frame carries individual acknowledgements for each forward packet so that only those frames that were lost or corrupted need to be re-sent. Selective repeat ARQ has higher asymptotic throughput than stop-and-wait ARQ at the cost of a longer cycle time.

Precise frame timing. When the precise timing of the start of a transmission is unknown, a synchronization preamble is sent at the start of the transmission to synchronize the sender and receiver. This sync preamble can be a large portion of the link turnaround time.

Precise synchronization of the transmission cycle permits use of a very short sync preamble on each transmission. However, this degree of precision in the timing of transmissions requires close coordination between the ARQ protocol and the modem.

Code combining modem. Code combining in the modem can be integrated with an ARQ protocol to dynamically adapt the FEC code rate on a packet-by-packet basis [3, 4]. FEC code bits are computed for each packet to be sent, but packets are initially sent at rate R=1. The remaining FEC bits are stored at the sending node. For example, if a rate 1/4 code is used, three additional k-bit sets of FEC bits will be stored for each k-bit packet sent.

For each packet received with errors (detected using a CRC), the soft decisions are stored and a retransmission is requested. Retransmissions carry alternating sets of the FEC bits for the packet, and the soft decisions from all receptions of a packet are combined by the channel decoder to improve its error correcting power. In this way, the number of sets of FEC bits sent for each packet is minimized, and can vary from packet to packet, resulting in packet-by-packet data rate adaptation with no more on-air overhead than is already necessary for ARQ operation.

Duplex client data flow. Most wireless links operate in only one direction at a time at the physical layer. Many wireless ARQ protocols also permit client data to flow in only one direction on the data link, with the return channel carrying only acknowledgements. A link-layer handshake is required to reverse the flow. However, it is certainly possible to support duplex flow of client data over a single logical link by including both client data and ARQ acks in each transmission.

3. DESIGN OF THE STUDY

A simulation study was designed to explore the interactions of these ARQ features, the wireless medium, and Internet applications.

Metric. We are interested in the interaction of latency in the wireless network with the higher-layer protocols, so we chose a time-based metric: the time to complete an application-layer session via the wireless channel.

Workload Factors. The workload is an email session that transfers a 5000-byte message. Our Internet-style workload uses SMTP and TCP. The sequence of transmissions for SMTP is shown at the bottom of the page.

To determine the impact of Internet-style workloads, we also simulated use of HMTP (a "wireless-friendly" version of SMTP) to transfer the message. HMTP exchanges the same messages as SMTP, but all of the client messages are sent in one transmission following the 220 message from the server. All of the remaining server messages are then sent in a single final transmission. TCP was not used in our wireless-friendly workload.

System Factors. Because the "precision frame timing" feature and the code combining modem both require integration of the modem with the ARQ protocol, they are considered as a single feature in this study. This leaves us with three system factors for our simulation study: ARQ mode, integrated modem, and duplex client data. A fractional factorial design [5] for these three factors is described in the table below. Each combination of features is similar to one of the existing HF ARQ protocols, and carries a corresponding designation.

Designation	**Selective Repeat**	**Integrated Modem**	**Duplex Client Data**
HDL-1d	no	**Yes**	**Yes**
HDL	**Yes**	**Yes**	No
5066-1s	no	no	No
5066	**Yes**	no	**Yes**

Environmental Factors. The channel used in this study is an HF radio skywave (ionospheric) channel, described by the Walnut Street model [6], which was validated [7] in 1997 as a component of the NetSim simulator by the Defense Information Systems Agency (DISA) Joint Interoperability Test Center (JITC).

This investigation used a multipath channel with independent Rayleigh fading of two equal-power paths separated in time by 2 ms with Doppler spreads of 1 Hz. The median SNR variation from its long-term average followed a lognormal distribution with a 10 second time constant.

The simulation study was performed for three instances of this channel:

- A stable, high-SNR channel with a fixed long-term average SNR of 25 dB, and 4 dB amplitude of the "mid-term" lognormal variation
- A stable, low-SNR channel with a fixed long-term average SNR of 5 dB, and 4 dB variation
- A transitional channel with average SNR dropping from 25 dB to 5 dB over a period of one hour, and 30 dB amplitude lognormal variation

Simulation Approach. Each simulation is run for at least an hour of simulated time, and the latency of all messages delivered is averaged.

The standalone modem is a MIL-STD-188-110B HF data modem, with data rates of 75 through 9600 bps. The integrated modem is the third-generation burst modem from MIL-STD-188-141B Appendix C which operates at 4800 bps. Simulation models for both modems use performance data measured on commercial implementations.

The simulated systems that use an external modem automatically adapt the data rate of that modem for each new transmission using the self-identifying preambles of the MIL-STD-188-110B modem. The initial data rate is set in these simulations using the user-specified SNR. Although this is unrealistic for real systems, it may be approximated in practice by measuring the channel SNR during automatic link establishment. The code-combining modems adapt data rate packet-by-packet within a transmission, and do not need a starting data rate.

Server	**Client (SMTP)**
`220 server.goodguys.gov ESMTP Sendmail ready`	
	`HELO myhost.mycompany.com`
`250 server.goodguys.gov`	
	`MAIL From:<myname@mycompany.com>`
`250 <myname@mycompany.com>... Sender ok`	
	`RCPT To:< recipient@somewhere.mil >`
`250 <recipient@somewhere.mil>... Recipient ok`	
	`DATA`
`354 Enter mail, end with "." on a line by itself`	
	`<5000 byte message goes here>` `.`
`250 Message accepted for delivery`	
	`QUIT`
`221 server.goodguys.gov closing connection`	

Two of the protocols simulated (5066 and HDL) correspond to existing HF ARQ protocols, and the simulation models produce performance consistent with commercial implementations of those protocols. The other two protocols are minor variations on existing protocol models:

- HDL-1d is a stop-and-wait, duplex version of the 3G HDL protocol, formed by sending a robust control packet optionally followed by a data packet (flagged in the control packet) in each transmission.
- 5066-1s is derived from 5066 by forcing each frame to carry only one data packet (stop-and-wait ARQ). Simplex operation results from using only Data or Ack packets (not Data+Ack packets).

In all cases, the ARQ protocols implemented a simple filtering mechanism to reduce the impact of inappropriate TCP retransmissions: each TCP segment submitted to the ARQ entity for transmission was compared to queued segments and discarded if an earlier copy was already awaiting or in transmission.

4. RESULTS

The average message latencies for the SMTP/TCP Internet-style workload are shown in Table 1, with the corresponding message throughputs graphed in Figure 1. Note the rising trend in throughput among the first three protocols at high SNR versus the falling trend at low SNR.

Table 1: Average Message Latency for SMTP (s)

Designation	25 dB SNR	5 dB SNR	Transition
HDL	206	343	242
HDL-1d	161	474	231
5066	119	667	290
5066-1s	321	2000	2000

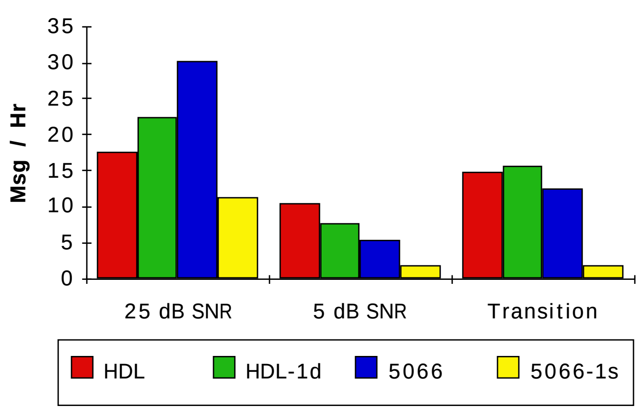

Figure 1: Message Throughput for SMTP

Table 2 and Figure 2 show the results for the wireless-friendly workload using HMTP without TCP. The performance boost compared to the SMTP/TCP case is substantial. Another notable change is the drop in relative performance of the HDL-1d protocol. A thorough analysis of these results is presented in the next section.

Table 2: Average Message Latency for HMTP (s)

Designation	25 dB SNR	5 dB SNR	Transition
HDL	57	110	70
HDL-1d	92	232	126
5066	33	439	92
5066-1s	120	507	735

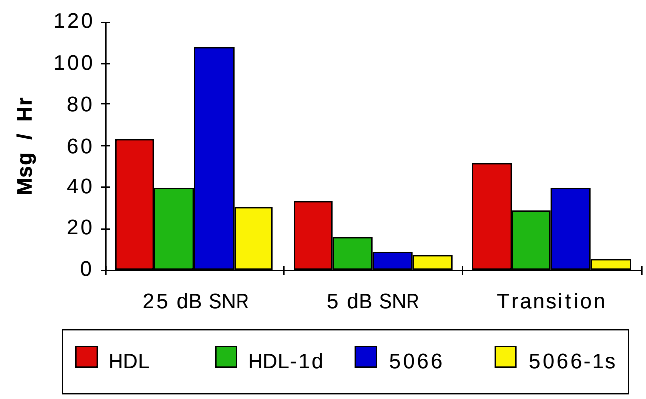

Figure 2: Message Throughput for HMTP

5. ANALYSIS

These simulation results can be analyzed to yield performance effects of our three ARQ features from three different perspectives:

- the impact of each feature on SMTP/TCP throughput
- the interaction of the features with channel variability
- the interaction of the features with the different workloads (i.e., Internet-style versus wireless-friendly)

In each case, the magnitude of the effect due to each feature is computed from the simulation results, along with some of the interactions among effects. (Because this is a fractional factorial experiment, some effects are confounded with high-order interactions.) The importance of each feature is gauged by the fraction of overall performance variation that its effect represents.

ARQ Effects on SMTP Performance. Beginning with the case of SMTP/TCP over a channel with stable SNR (high or low), the expected large positive effect due to high SNR dominates all other effects (see Table 3).

Table 3: ARQ Effects for SMTP/TCP (All SNR)

Feature	Effect (% of overall variation)	Sign
High SNR	63 %	+
Duplex Client Data	12 %	+
Selective Repeat	9 %	+
Integrated Modem	2 %	+
Interactions		
SNR * Duplex Data	11 %	+
SNR * Integr Modem	3 %	–
SNR * SelectRpt	1 %	+

The effects of our features at high SNR (Table 4) show that support for duplex client data is highly beneficial in supporting Internet-style applications. Selective repeat is also of some value, despite the prevalence of short messages. However, the integrated, code-combining modem had negligible effect for Internet applications at high SNR.

Table 4: ARQ Effects for SMTP/TCP (High SNR)

Feature	Effect %	Sign
Duplex Client Data	74 %	+
Selective Repeat	26 %	+
Integrated Modem	0 %	–

Examination of the simulation log files shows that at high SNR both modems transfer data quickly and efficiently, so message latency is strongly dependent on protocol overhead. The SMTP handshakes and TCP acks in this Internet-style workload require frequent changes of direction of client data flow. The simplex protocols must exchange link-layer control packets for every reversal of client data flow, resulting in noticeable performance loss. The duplex protocols avoid this overhead, and further boost TCP performance by delivering TCP acks for completed segments while subsequent segments are in progress.

At low SNR, however, we find a remarkable turnaround: use of the integrated, code-combining modem is the dominant effect, while support for duplex client data is virtually irrelevant (see Table 5).

Table 5: ARQ Effects for SMTP/TCP (Low SNR)

Feature	Effect %	Sign
Integrated Modem	74 %	+
Selective Repeat	26 %	+
Duplex Client Data	0 %	+

The benefit of the integrated modem at low SNR flows from two sources: the error rate benefit from code combining, and the packet-by-packet adaptation of the effective data rate.

- The protocols using a stand-alone modem must select a single data rate for each transmission; at 5 dB SNR, 300 bps results in a packet error rate of over 50%, while 150 bps is nearly error-free. The adaptive data rate algorithm tends to switch between these two rates, achieving a sustained throughput of around 100 bps.
- Despite using a 4800 bps waveform at 5 dB SNR, the HDL protocol typically requires only five transmissions of an ARQ packet for error-free operation after code combining. The resulting sustained throughput is over 400 bps (consistent with results in [4]).

Of course, these sustained throughputs occur only during transmission of long messages. Link turnarounds, TCP acks, and short SMTP handshake messages dilute the latency advantage of the integrated modem.

Interaction of ARQ Features with Channel Variability. The effect of high variability in channel SNR can be gauged by comparing the performance of each test case in the transitional channel (25 down to 5 dB SNR with 30 dB variation) to the mean of the corresponding results obtained in the two stable-SNR channels. The results in Table 6 show that each of our three features has some value. This is consistent with Table 1 and Figure 1, where we see that the 5066-1s protocol, which has none of the features, achieves much lower performance than the other three protocols, each of which has two of the three features.

Table 6: ARQ Effects for SMTP/TCP (Transition)

Feature	Effect	Sign
Duplex Client Data	34 %	+
Integrated Modem	27 %	+
Selective Repeat	25 %	+
Stable SNR	5 %	+
Interactions		
StableSNR*IntModem	8 %	–
others	0 %	+

The negative interaction between Stable SNR and the Integrated Modem in Table 6 indicates that the integrated modem loses less performance when SNR is not stable.

Interaction of ARQ Features with Workload. Comparing Figures 1 and 2, it is apparent that use of a “wireless friendly” e-mail application results in roughly threefold performance improvement with any of our ARQ protocols. Tables 7 and 8 present in more detail the effects and interactions of workload and ARQ features.

Table 7: ARQ / Workload Interaction (High SNR)

Feature	Effect %	Sign
Wireless Workload	45 %	+
Selective Repeat	24 %	+
Duplex Client Data	11 %	+
Integrated Modem	2 %	–
Interactions		
Workload * SelectRpt	14 %	+
Workload * IntModem	2 %	–
Workload * Duplex	2 %	+

As noted previously, performance in a high-SNR channel is largely determined by protocol overhead rather than time to deliver the client data. The wireless workload eliminates most link turnarounds, which directly reduces latency. The resulting long transmissions take advantage of selective repeat ARQ operation, so we also find a large benefit from use of Selective Repeat ARQ and a strong positive interaction between these two individual effects.

At low SNR, however, use of an integrated modem was again the dominant effect, more important even than use of a wireless-friendly workload. Both the precision timing of the integrated modem and the processing gain of code-combining contribute to reducing message latency in this challenging channel.

Table 8: ARQ / Workload Interaction (Low SNR)

Feature	Effect %	Sign
Integrated Modem	37 %	+
Wireless Workload	28 %	+
Selective Repeat	12 %	+
Duplex Client Data	5 %	–
Interactions		
Workload * IntModem	9 %	+
Workload * Duplex	6 %	–
Workload * SelectRpt	3 %	+

6. CONCLUSIONS

Each of the three features studied here – selective repeat operation, duplex client data, and an integrated code-combining modem – improved message delivery performance in some circumstances.

- Selective repeat ARQ operation provided some benefit under all circumstances, and was the dominant factor for the wireless-friendly workload in the high SNR channel.
- When SNR is high, protocol overhead dominates the time to deliver client data over the channel. Unless the ARQ protocol supports duplex flow of client data, the numerous short messages of Internet applications (SMTP in this study) and the inherently duplex nature of TCP cause frequent, time-consuming ARQ handshakes to reverse the direction of client data flow.
- In our low-SNR channel, use of the integrated modem dominated all other factors, including the choice of Internet-style or wireless-friendly client protocols.

These results suggest the following general conclusions for supporting Internet-style applications over wireless subnetworks in general:

- When the channel supports high data rates, protocol efficiency dominates. Internet-style applications then need support for duplex client data flow. Selective repeat operation is also beneficial.
- In low-SNR channels, precise handshake timing reduces protocol breakdowns and a code-combining modem wrings maximum benefit from all received energy. The other ARQ features are less important.

7. FUTURE WORK

None of the protocols simulated here included all three of the features investigated. It would be interesting to see whether inclusion of all three features would provide superior “all weather” performance.

REFERENCES

1. H. Balakrishnan, *et al.*, “A Comparison of Mechanisms for Improving TCP Performance over Wireless Links,” *IEEE / ACM Transactions on Networking*, **6**(5):756-769, 1997.
2. H. Chaskar and U. Madhow, “TCP over wireless with link level error control: Analysis and design methodology,” *Proceedings of MILCOM ’96*, IEEE, 1996.
3. D. Chase, “Code Combining – A Maximum Likelihood Decoding Approach for Combining an Arbitrary Number of Noisy Packets,” *IEEE Trans Comm*, **33**(4):385-393, 1985.
4. M. Chamberlain, *et al*., “Performance of U.S. MIL-STD-188-141B Appendix C Data link protocols,” *Proceedings of Eighth international conference on HF radio systems and techniques*, IEE, 2000.
5. R. Jain, *The Art of Computer Performance Evaluation*, (chapters 16 and 19), Wiley, New York, 1991.
6. E. Johnson, “Interactions Among Ionospheric Propagation, HF Modems, and Data Protocols,” *Proceedings of Tenth Ionospheric Effects Symposium*, NTIC, 2002.
7. R. Bullock, “SCOPE Command High Frequency (HF) Network Simulation Model (NetSim-SC) Verification and Validation Report,” Joint Interoperability Test Command, Ft. Huachuca, AZ, January 1997.