An Asymmetric, Dynamic, Energy-conserving ARQ Protocol

Jason Redi^a, Chiara Petrioli^b, Imrich Chlamtac^c

^{a} BBN Technologies, GTE Internetworking, Cambridge, MA, USA, *redi@bbn.com*
^{b} Politecnico di Milano, Italy, *Chiara.Petrioli@elet.polimi.it*
^{c} The University of Texas at Dallas, TX, USA, *chlamtac@utdallas.edu*

Abstract

Mobile computing implies reliance on a portable, self-contained power source and increased difficulty of communications due to the highly-variable radio channel. In order to ensure reliable, verified, error-free delivery of data in the presence of multipath fading, automatic repeat request (ARQ) schemes are used to synchronize and acknowledge the transmission of data between the base station and the mobile node. In this paper, we build upon previous designs of energy-conserving variants of the classical Selective Repeat protocol to design a dynamic, asymmetric energy-conserving protocol. Primarily, in this work, we move the complexity of the protocol to the base station, which can optimally assign on a packet-by-packet basis the set of energy-conserving ARQ protocol and parameters which is best suited to the current radio channel, traffic and energy requirements. In this manner, the best delay vs. energy trade-off of previously proposed protocols can be incorporated in a practical, dynamic way. Through simulation we evaluate the average packet delay of different classes of service under strict energy constraints and varying channel characteristics. The results derived here show the significant advantage of quickly adapting the protocol to changing channel fading characteristics so as to maintain a required energy vs. delay trade-off between multiple priority classes of traffic.

1 INTRODUCTION

Mobile wireless devices carry the burden of an energy-constraint, not typically found in wired networks. In order to be easily mobile, a wireless device must rely on a small portable power source in the form of a battery. Given that current battery capacities are not expected to increase in potential more than 30% in the near future [1], and the wireless network transceiver can typically use 15-30% of the power of a typical mobile computer [2] (with the display the only part using more power), it is of great importance to include energy-conservation in the design of wireless network protocols.

To insure the reliable delivery of packets in the error-prone wireless channel, automatic repeat request (ARQ) protocols are employed to acknowledge correct packet reception. Since transmission typically uses twice as much energy as reception, traditional ARQ protocols are extremely energy-intensive as they specify a single acknowledgement for every correctly received packet. In a previous paper, the authors introduced energy-conserving variants of the classical ARQ protocols [3], which were designed to reduce the number of acknowledgements sent over the uplink, and therefore substantially reduce the energy required by the ARQ protocol, while maintaining an average packet delay which is within normal application constraints. The principle behind the proposed protocols is intuitive, allowing multiple outstanding packets to be sent without acknowledgment by the base station, and then have the mobile node acknowledge a group of packets with just a single acknowledgment. In this paper we introduce a practical extension to this protocol to show how it can be used as an asymmetric ARQ energy-conserving protocol which can dynamically adapt to changing channel properties while balancing the trade-off of energy vs. delay for different types of wireless traffic.

1.1 The utility of asymmetry

Wireless data services are experiencing the same convergence of traffic types as the rest of the networking community. The wireless data networks of the future will be expected to transport a mix of video, audio, and asynchronous data to a variety of devices which will have a wide range of capabilities in terms of buffering, allowable complexity, transmitter power, disconnection tolerance, etc. It is therefore important that future wireless network protocols be flexible enough to support traffic convergence and end-station variety.

In order to more easily support a multitude of traffic and nodal requirements while still maintaining fair access across the entire cell we propose moving the majority of the decisions regarding when acknowledgments should be sent (therefore the complexity of the protocol) to the base station. In this model, base station specifies when and what kind of ARQ replies the mobile node is to give.

Asymmetric energy-conserving ARQ protocols which dynamically adapt to changing conditions have numerous advantages:

Traffic Management: Varying traffic patterns and QoS requirements can be more easily supported since the base station has current knowledge of all arrival processes.

Mobile Power Management: Traffic for nodes with a very limited amount of available reception energy can be given a higher priority. Additionally, energy-efficiency of the ARQ protocol running at each node can be more effectively managed by the base station by including all nodes' battery capacities and expected error rates.

Cost: The cost of the base station is, in effect, shared across all the network subscribers. Protocol changes, cell bandwidth management, and other complex operations are therefore cheaper at the base station than at the mobile nodes.

Device independence: The available power, buffering, and other transmission capabilities can vary greatly across the range of devices. Placing the decisions and bulk of the protocol at the base station allows for simpler and a more uniform interface.

The use of an asymmetric channel in a wireless data network has been previously described in [4]. They proposed a new link-level wireless protocol called AIRMAIL which is asymmetric and includes acknowledgments for multiple packets for the purpose of energy conservation. However, this work focuses mainly on the reduction of the protocol complexity at the mobile node in terms of network code size, and does not specifically investigate the use of multiple packet priorities, nor evaluate the energy vs. delay trade-off. The General Packet Radio Service (GPRS) in GSM utilizes a selective repeat method of acknowledging transmitted "bursts" [5]. The transmission of all 57 bit bursts which comprise a single complete, 1600 byte maximum, link layer packet must be successful before subsequent bursts are sent. GPRS therefore only selectively acknowledges the bursts that comprise a complete link layer packet such that there is no concatenation of the transmission of multiple link layer packets. Furthermore, there is no provision for adapting the ARQ scheme to the transmission characteristics or other factors.

In the following two sections, we first introduce the system model as well as the assumptions we use for simulation. We then overview the energy-conserving ARQ protocols, and explain how they can be used to build dynamic, asymmetric protocols. We follow with simulation results that show the utility of the ideas and discuss their advantages.

2 SYSTEM MODEL AND PROTOCOL DESCRIPTION

We consider a system where a base station communicates with a mobile node through a radio channel of bandwidth B. The model is easily generalized to represent a multi-station environment each with multiple nodes, by replicating, without restriction, the single base station model. We assume the communication to be packet-oriented, the time to be slotted and the base station's transmissions to be synchronized to the beginnings of slots. The packet length c is constant, and exactly one packet can be transmitted during one slot. To mitigate the adverse effects on performance due to time-correlated multipath fading while limiting the energy used for uplink communication, the mobile node acknowledges the packet transmission status (correct/incorrect) and the base station retransmits the lost packets according to an *energy-conserving Selective Repeat ARQ protocol*. For this protocol, the mobile node conserves energy by acknowledging groups of packets sent by the base station with a single acknowledgment. The acknowledgments follow a selective repeat style such that they include a bit-map denoting the reception status of all the packets transmitted since the last ACK. The base station therefore only retransmits the incorrectly received packets.

In [3] we introduced two types of Selective Repeat-based Energy-Conserving ARQ protocols. In the first type, described as *Windowed Feedback with Selective Repeat*, acknowledgments are sent in response to a group (or "window") of W packets, a duplicate packet reception, or the occurrence of a time-out. Since packet buffer space at the base station should not be freed until it is known that the packet was received error-free at the mobile node, these time-outs are set to occur when the time between two ACKs exceeds a fixed threshold t, either because of extremely light traffic or extremely bad channel conditions. For this protocol, the transmitter acts according to the following rules:

- It sends packets in order, as long as they belong to the current window. When a window has been completely transmitted or a time-out occurs, the most recently transmitted packet is continuously retransmitted until an ACK is received.
- Upon reception of an ACK, the buffers associated to the correctly received packets are freed. The transmitter window is then updated to the next W unacknowledged packets.

For the second type of protocol, described as *Instantaneous Feedback with Selective Repeat*, both the mobile node and the base station act as in the Windowed Feedback case. However, the mobile node also uses negative acknowledgments to reply to out-of-order packet receptions. The reception of an ACK triggers both the retransmission of the erroneously received packets as well as an update of the transmitter window.

Beginning with these two types of protocols, we can build an energy-conserving ARQ scheme to be largely the responsibility of the base station. It is here that information regarding the priority of packets and the type of data is known, as well as information regarding available received transmission power and error rates for the individual nodes. Using this information, the base station can quickly react to dramatic changes in the channel capabilities or traffic load, and reflect these in the ARQ protocol. Similar to a base station dynamically changing the node's wakeup schedules to suit the traffic load [6], an *asymmetric, dynamic energy-conserving ARQ protocol* allows the base station to use just a few bits in each packet to request or decline acknowledgments from the nodes to suit the channel error process, packet priorities, and node's available transmitter power.

This new type of protocol is asymmetric since the base station regulates the transmission of acknowledgments. The differences shown between the windowed and instantaneous negative acknowledgments can be incorporated into a general scheme in which three extra bits in the packet header act to query the mobile node for its acknowledgments.

These bits can be used for the following:

Request for acknowledgment bit: Requests an acknowledgment to be sent in response to this packet. As before, an acknowledgment explicitly states which packets were in error. This bit can be used at the end of a window of unacknowledged frames to result in windowed feedback, or be used for supporting different delay vs. energy trade-offs across different types of traffic.

Instant negative mode bit: Set the mode of the mobile node to send negative acknowledgments when errors occur in this or future packets.

Selective or Go-Back-N mode bit: Switch between the two different types of packet buffering. In this paper, we only consider a selective repeat acknowledgement scheme since it has better delay vs. energy performance and makes more efficient use of the wireless channel. However, Go-Back-N acknowledgements may be appropriate for some applications or devices, such as the situation when the higher buffer requirements of selective repeat are difficult to meet.

Optimization of the acknowledgment scheme based on the power and error process can be performed with theoretical models of these protocols along with practical experiments. In general, a base station initiated acknowledgment protocol provides a large number of capabilities and performance improvements which take advantage of the inherently centralized architecture of the wireless cell.

3 PERFORMANCE AND DISCUSSION

To provide an example of protocol performance in a practical situation, we have simulated a system composed of a single base station attempting to deliver two types of packets to a mobile node. The packets are classified as either high-priority control packets or normal-priority data packets such that the base station always attempts to transmit a waiting control channel packet before a waiting data channel packet. We keep the ARQ protocol for each type of packet separate such that we can visualize the system as containing two virtual channels. The use and desirability of control channel is well known throughout the literature. A control channel typically will have a much lower data rate but carries information vital to the correct functioning of the network such as hand-off, bandwidth-provisioning, or emergency information. Therefore, we will be interested in maintaining a much lower average packet delay of control packets at the cost of a higher energy consumption. The trade-off can also be seen as the reduction of energy at the mobile node versus the timely receipt and verification of important information. The limitation of two priorities is merely for clarity of the discussion as it would be possible to use any number of different service types in this protocol.

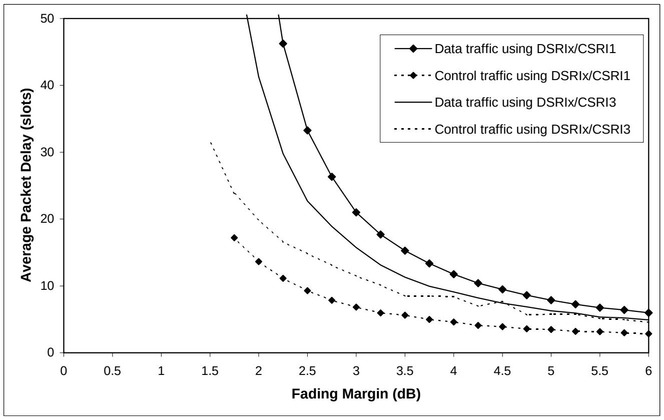

Figure 1: Average Delay vs. Fading Margin for two systems in slow fading; energy limited to 30%. The first system utilizes a variable instantaneous selective repeat data window size (DSRIx) and a fixed control window size of one (CSRI1). The second system utilized a variable data window size (DSRIx), but an instantaneous selective repeat control window size of three (CSRI3).

A total of 500,000 packets arriving according to a Poisson process with an average rate of $\lambda = 0.3$ packets/slot were delivered for each simulation run. 10% of the packets were randomly chosen to be the control packets, while the remaining were considered data packets. For each simulation run, we used an instantaneous selective repeat protocol with a window size of 1 or 3 for the control traffic (denoted as CSRI1 or CSRI3), and either a windowed or instantaneous selective repeat protocol with a variable window size for the data traffic channel (denoted DSRBx or DSRIx). To reduce coupling between the protocols which would complicate the evaluation, acknowledgements for the different types of traffic were considered separate, such that an acknowledgement for one type of traffic did not acknowledge any packets for the other type of traffic. The small control window channel size was selected due to the channel's low traffic load, along with its time-critical nature. For the data channel, we selected the ARQ window size to give the lowest delay while keeping energy under a threshold of $e_{\max}$ as measured by the average number of acknowledgements per slot. All acknowledgments are considered to use the same amount of energy for transmission. As pointed out in [3], long bursts of errors on the channel carrying the acknowledgement packets can increase the average delay of the protocol due to the sender believing that the packets, and not the acknowledgements, were being lost. For the simulations considered, we evaluated systems with $e_{\max} = 0.30$ and $e_{\max} = 0.10$. This deliberately varying selection of the data channel window size in order to optimize a certain function, despite changes in fading speeds and fading margins, exemplifies the utility of an asymmetric protocol: as we detect the change in error rate, we can dynamically optimize the window size while maintaining our other constraints. We can also

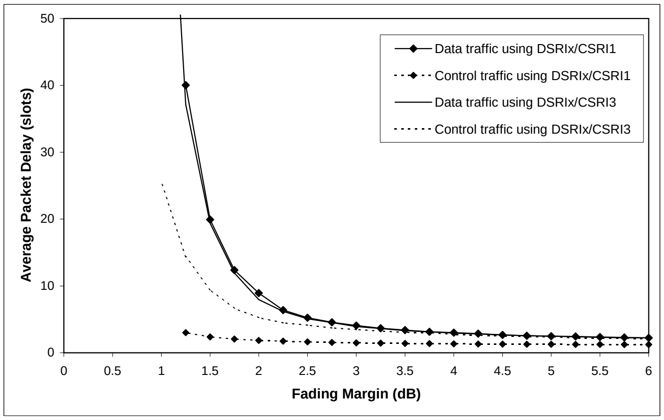

Figure 2: Average Delay vs. Fading Margin for two systems in fast fading; energy limited to 30%. The first system utilizes a variable instantaneous selective repeat data window size (DSRIx) and a fixed control window size of one (CSRI1). The second system utilized a variable data window size (DSRIx), but an instantaneous selective repeat control window size of three (CSRI3).

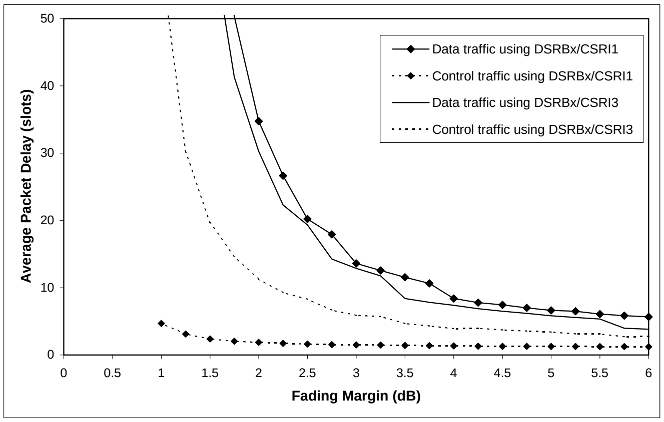

Figure 3: Average Delay vs. Fading Margin for two systems in fast fading; energy limited to 10%. The first system utilizes a variable buffered selective repeat data window size (DSRBx) and a fixed control window size of one (CSRI1). The second system utilized a variable data window size (DSRIx), but an instantaneous selective repeat control window size of three (CSRI3).

note the utility of using an energy-conserving ARQ scheme. The traditional stop-and-wait ARQ protocol would have resulted in an average energy usage of 0.30 in an error-free channel. Using this system, we can not only bound the energy to 0.30 even in the error-prone wireless channel, but also support much lower energy bounds such as 0.10.

We model both the downlink and uplink channels as Gilbert channels [7]. The ability of this model to capture the behavior of a narrowband fading channel, independently of modulation, coding techniques or packet length, was assessed in [8]. The pattern of feedback and packet errors is then described by two independent first-order Markov models. Two states (*correct* and *incorrect*) then represent the status of the channel during the current slot. The transition probabilities can be easily computed and only depend on the normalized Doppler frequency and the steady-state packet error rate.

Figures 1 through 3 depict the average packet delay versus fading margin for each type of traffic using the asymmetric ARQ protocol. Delay is defined as the number of slots between the arrival of the packet at the base station and the time when the packet can be sent to the higher levels of the protocol stack at the destination node (i.e. when all preceding packets have also been successfully received). Each plot contains two pairs of lines, each corresponding from the same simulation parameters. The fading margin (F), as plotted along the x-axis of each figure, is effectively the signal to noise ratio and is related to the steady-state packet error rate (ϵ) by $\epsilon = 1 - e^{1/F}$. A high fading margin therefore represents a low steady state error rate, independent of the mobile's speed. Two of the figures (2 and 3) depict performance with a "fast" normalized Doppler frequency ($f_D NT$) of 1.0 Hz, while Figure 1 depicts performance with "slow" normalized Doppler frequency of 0.02 Hz. All else being equal, a slower Doppler frequency produces longer, though less frequent bursts of errors, while with a fast Doppler frequency, the error process becomes time-independent across packets.

For clarity, we have not plotted all combinations of data and control channel acknowledgement schemes and instead plotted only those which achieved the best delay performance for the energy constraints. For Figures 1 and 2, due to the limited amount of feedback, buffered selective repeat acknowledgements in the data channel resulted in average delays up to 20 slots greater than the instantaneous selective repeat protocols. However, using instantaneous acknowledgements requires a higher lower bound on energy that is tightly coupled to the channel state transitions. For this reason Figure 3 does not contain the instantaneous acknowledgements as they are bounded to a fading margin of 3.25 dB or above with the same performance as the buffered acknowledgements.

By observing Figure 1, we can see that the selection of control window size has an effect not only on the control packet average delay, but on the data packet delay as well. As would be expected, requiring fewer acknowledgements in the high priority traffic allows the base station to deliver data packets with fewer delays due to retransmissions of control packets due to lost acknowledgments. However, this characteristic is not observable in fast fading (Figure 2). Due to the shorter bursts of errors in fast fading, there is a much lower probability of missing acknowledgements in the control packet window holding up the delivery of data packets for extended periods of time. Fast fading also results in a lower average delay for all the traffic in the system. Although the steady state packet error rate is the same for equal fading margins, slower fading results in longer periods of unnecessary retransmissions due to extended bursts of errors in acknowledgement transmission.

Comparing Figure 1 and Figure 2 shows that the protocols are capable of meeting the 30% energy constraints for a smaller fading margin in fast fading than in slow fading. This is again due to longer bursts of errors on the uplink which limits the protocol ability to meet the energy constraints due to many more retransmissions of acknowledgements. Figure 3 reflects the ability of the buffered selective repeat protocols to meet energy constraints below those of the instantaneous acknowledgements. As mentioned, the system using instantaneous acknowledgements on the data channel was limited to delivering packets only down to fading margins 3.25dB or above. Even with the tight energy constraint of 10%, we can see that delivery of control channel information is still kept to a minimum while data channel traffic only experiences modest delays as we get below 2dB. For mobile nodes with low energy availability, using buffered selective repeat on the data channel can allow tremendous energy savings for acknowledgements while not adversely affecting control channel information. The cost of this protocol is in a slightly higher buffering requirement at the mobile node.

4 CONCLUSIONS

In this paper we have introduced and evaluated an asymmetric, dynamic, energy-conserving ARQ protocol, which is designed to adapt to changes in the radio channel, traffic and energy-requirements, in order to optimize the delay vs. energy trade-off. The complexity of the protocol is moved to the base station, which dynamically sets the protocol parameters, also providing support to different traffic classes.

Simulation results under varying traffic and channel conditions for both data and control packets allowed us to substantiate the need for a dynamic approach, and proved the ability of the proposed solution to achieve high energy saving while only slightly increasing the access delay.

References

[1] Powers, R.A., “Batteries for low power electronics,” *Proceedings of the IEEE*, 83(4):687–93, April 1995.

[2] Harris, E.P., Depp, S.W., Pence,E., Kirkpatrick, S., Sri-Jayantha, M., and Troutman, R.R., “Technology directions for portable computers,” *Proceedings of the IEEE*, 83(4):636–58, April 1995.

[3] Chlamtac, I, Petrioli, C. and Redi, J., “Energy-conserving Selective Repeat ARQ Protocols for Wireless Data Networks,” in *Proceedings of PIMRC ’98*.

[4] Ayanoglu, E., Paul, S., LaPorta, T.F., Sabnani, K.K., and Gitlin, R.D., “Airmail: a link-layer protocol for wireless networks,” *ACM/Baltzer Wireless Networks*, vol. 1, no. 1, pp. 47–60, 1995.

[5] Cai, J. and Goodman, D., “General Packet Radio Service in GSM”, *IEEE Communications*, pp. 122–131, October 1997.

[6] Chlamtac, I, Petrioli, C. and Redi, J., “Energy-conserving Protocols for Wireless Identification Networks,” scheduled for publication in *IEEE/ACM Transactions on Networking*, February 1999.

[7] Gilbert, E. N., “Capacity of a burst-noise channel,” *Bell Systems Technical Journal*, vol. 39, pp. 1253–1266, September 1960.

[8] Zorzi, M., Rao, R.R., and Milstein, L.B., “ARQ error control for fading mobile radio channels,” *IEEE Transactions on Vehicular Technology*, vol. 46, pp. 445–55, May 1997.