TSG-RAN Working Group 1 **_TSGR1#18(01)0124_**
Boston
January 15-19

Agenda item: AH24, HSDPA

Source: Lucent Technologies.

Title: Text Proposal for the TR 25.848

Document for: Discussion

1 Introduction

This contribution provides text proposal for HARQ (section 5.2) and FCSS (section 5.3) in HSDPA.

2 Text Proposal

5.2 Hybrid ARQ (H-ARQ)

(WG1 note: the contents of this section are for further review and alignment with WG2)

H-ARQ is an implicit link adaptation technique. Whereas, in AMC explicit C/I measurements or similar measurements are used to set the modulation and coding format, in H-ARQ, link layer acknowledgements are used for re-transmission decisions. There are many schemes for implementing H-ARQ - Chase combining, Rate compatible Punctured Turbo codes and Incremental Redundancy. Incremental redundancy or H-ARQ-type-II is another implementation of the H-ARQ technique wherein instead of sending simple repeats of the entire coded packet, additional redundant information is incrementally transmitted if the decoding fails on the first attempt.

H-ARQ-type-III also belongs to the class of incremental redundancy ARQ schemes. However, with H-ARQ-type-III, each retransmission is self-decodable which is not the case with H-ARQ-type II. Chase combining (also called H-ARQ-type-III with one redundancy version) involves the retransmission by the transmitter of the same coded data packet. The decoder at the receiver combines these multiple copies of the transmitted packet weighted by the received SNR. Diversity (time) gain is thus obtained. In the H-ARQ-type-III with multiple redundancy version different puncture bits are used in each retransmission.

AMC by itself does provide some flexibility to choose an appropriate MCS for the channel conditions based on measurements either based on UE measurement reports or network determined. However, an accurate measurement is required and there is an effect of delay. Also, an ARQ mechanism is still required. H-ARQ autonomously adapts to the instantaneous channel conditions and is insensitive to the measurement error and delay. Combining AMC with H-ARQ leads to the best of both worlds - AMC provides the coarse data rate selection, while H-ARQ provides for fine data rate adjustment based on channel conditions.

The choice of H-ARQ mechanism however is important. There are two main ARQ mechanisms - selective repeat (SR) and stop-and-wait (SAW). In SR, only erroneous blocks are re-transmitted. A sequence number is required to identify the block. Typically, in order to fully utilize the available channel capacity the SR ARQ transmitter needs to send a number of blocks while awaiting a response (or lack of it in this case). Hence when combined with H-ARQ the mobile needs to store soft samples for each partially received block. Thus mobile memory requirements can be huge. More importantly, H-ARQ requires that the receiver must know the sequence number prior to combining separate re-transmissions. The sequence number must be encoded separately from the data and must be very reliable to overcome whatever errors the channel conditions have induced in the data. Hence a strong block code is needed to encode the sequence information - increasing the bandwidth required for signaling.

While Stop-and-Wait (SAW) is a simple form of ARQ, it could result in inefficiency because during the time that the transmitter is waiting for an acknowledgement (or negative acknowledgement) from the receiver (feedback delay), the channel is unused. This problem is alleviated by a multi-channel version of SAW, called N-channel SAW [1][2] wherein the transmitter can continue sending new data blocks to the user while awaiting feedback for a previously transmitted block. Of course, channel utilisation will improve for SAW when there are multiple users present. However, in a multi-user context, N-channel SAW provides the additional flexibility to schedule the same user in successive slots to take advantage of that user's good channel conditions. The number of channels, N, can be determined based on the feedback delay. For example, if the feedback delay is (N-1) TTIs, then N parallel channels can be used.

5.2.1 Timing in HARQ Operation

N-channel SAW can be operated in a variety of modes. In synchronous mode, the transmitter has flexibility in choosing the time for the first transmission of a data block. Retransmissions of the data block can take place only at fixed time intervals relative to the first transmission. If the first transmission of a data block takes place in TTI #i, then retransmission to the user (if any) must take place in TTI number (i+N) and if that retransmission fails, then the next retransmission will take place in TTI number (i+2N) and so on. A variant of this approach is to allow partially asynchronous operation, wherein retransmission can take place at an interval that is at any integer multiple of N. As in the previous example, suppose the first transmission to the user takes place in TTI number i, then a retransmission can take place in any subsequent TTI number i+jN where j is an integer. This affords the transmitter some flexibility in assigning TTIs corresponding to that channel to other users. An example of the partially asynchronous approach is given in Figure 1~~Figure 1~~ where the retransmission to user A is pre-empted to accommodate user B's new transmission and retransmission.

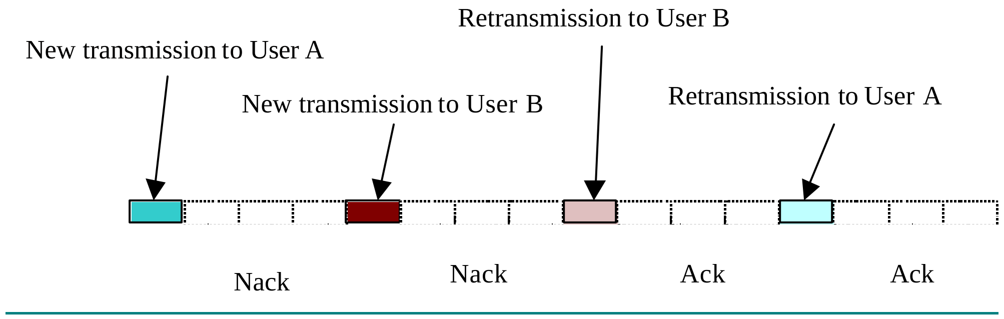

Figure 1~~1~~. An example of "partial asynchronous" operation

A greater degree of flexibility can be obtained by allowing fully asynchronous operation wherein transmissions/retransmissions to any user on any channel can be done at any time [3][4]. This allows full realisation of multi-user diversity gains. While synchronous and partially asynchronous modes do not require any overhead bits, fully asynchronous operation does require that a channel identifier be transmitted to the user, thus requiring $\log_2 N$ overhead bits. For typical code block sizes and number of multiple channels supported on the HS-DSCH, this would represent a small percentage overhead. The choice of synchronous, partially asynchronous or fully asynchronous operation of the N-channel SAW should be based on performance, flexibility and complexity.

An example of fully asynchronous operation for the same case as in Figure 1~~Figure 1~~ is depicted in Figure 2~~Figure 2~~. The retransmission to user A can now be performed in the 5th TTI after the original transmission even though the interval between transmissions is not an integer multiple of N. In case of "partial asynchronous" operation, the retransmission to user A has to be delayed to 12th TTI after the original tranmsission. Therefore, the time to complete transmission in a partial asynchronous scheme could be larger compared to a full asynchronous scheme. Moreover, with fully asynchronous operation, a large amount of data can be sent to a user when the channel conditions are favorable (Figure 2~~Figure 2~~) thus fully exploiting the multi-user diversity gains.

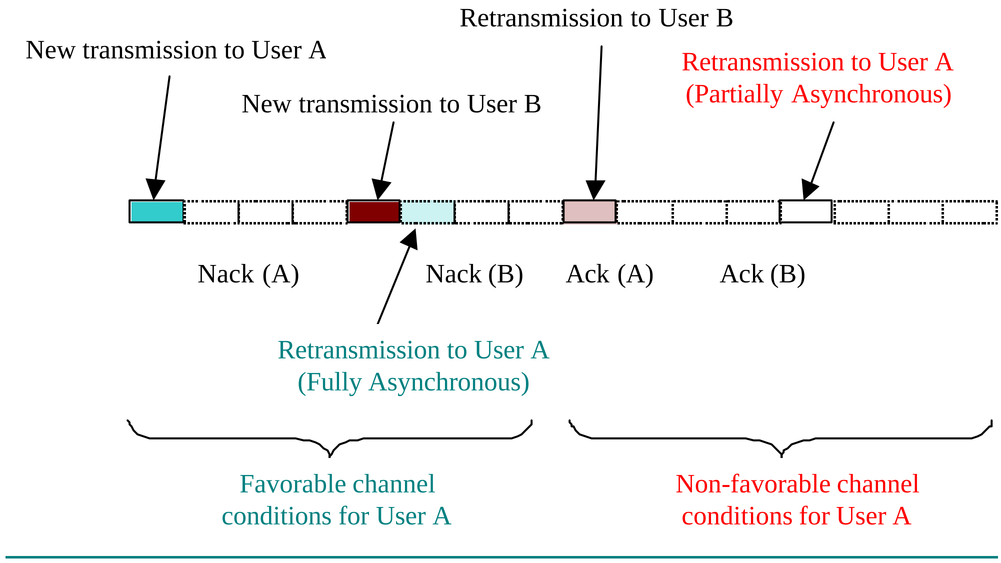

Figure 22. An example of fully asynchronous operation

If the feedback delay from the UE to Node B is deterministic, then no indication of the channel being acknowledged is necessary on the uplink for any of these schemes.

5.2.2 Interaction of HARQ and AMC

H-ARQ autonomously adapts to the instantaneous channel conditions and is insensitive to the measurement error and delay. Combining AMC with H-ARQ leads to the best of both worlds - AMC provides the coarse data rate selection, while H-ARQ provides for fine data rate adjustment based on channel conditions. Another interaction between HARQ and AMC is in the choice of MCS for retransmission. The choice of MCS for retransmission could be the same as that of the original transmission (fixed) or could be allowed to change (adaptive).

The power available for HS-DSCH is continuously changing (on a slot-by-slot basis) due to variations in the power used by power-controlled circuit switched users. The C/I seen by a user is also varying due to varying interference from neighboring cells and/or changes in channel quality due to fading etc. Furthermore, with asynchronous HARQ operation, the time between the transmissions/retransmissions could be longer because a retransmission to a user can be pre-empted by a transmission/retransmission to another user. Therefore, it is likely that the channel conditions, available power and code space are different between transmissions/retransmissions that need to be HARQ combined. Under these conditions, non-adaptive schemes that do not allow change of MCS on retransmissions will have to abort transmission and that can result in degraded system performance. Note that even if the number of codes available for HS-DSCH changes comparatively slowly, a small change in the number of codes make it impossible to perform a retransmission at the same MCS for non-adaptive HARQ schemes.

An adaptive scheme provides the most flexibility in matching the MCS level to the channel not just for the first transmissions but also for retransmissions.

5.3 Fast Cell Selection (FCS)

With Fast Cell Selection, the UE does not receive simultaneous data transmission from multiple cells and therefore performs no combining of traffic channels carrying packet data. Instead, the UE selects the best cell every frame from which it requests the data to be transmitted. The uplink DPCH is used to indicate the required cell from which the network should direct its data transmission to the UE on a frame by frame basis. This technique is a very special case of Site Selection Diversity (SSDT) and applies only to the HS-DSCH. In the case of SSDT, each cell is assigned a temporary ID and UE periodically informs a primary cell ID to the connecting cells. The non-primary cells not selected by the UE switch off their transmitter. However, in the case of Fast Cell selection, the UE selects the best cell every frame from which it wants to receive data on the HS-DSCH. HS-DSCH data is then transmitted to the UE from this cell only.

With fast cell site selection (within a Node B or between Node Bs), the channel conditions, available power and the code space will likely be different in the new cell. With N-channel HARQ, it is quite probable that some of the code blocks will have pending recovery while FCS is performed. Therefore, a true HARQ scheme should be able to do IR/combining across transmissions/retransmissions from different cell sites. Moreover, if the MCS is constrained to be the same for re-transmissions from the newly selected cell, the originally used MCS and multi-code information will have to be signaled by the UE to the new cell. In contrast, with adaptive HARQ operation, the FCSS will be completely transparent, i.e. any MCS and the available code space from the new cell can be used for retransmissions.

3 References

[1] "Clarifications on Dual-Channel Stop-and-Wait HARQ", TSG-RAN #18(01) 0048, Motorola.

[2] "Considerations on HSDPA HARQ concepts", TSG-RAN #18(01) 0007, Nokia.

[3] "Asynchronous and Adaptive IR (A^2IR) for HSDPA", Lucent Technologies, TSG-RAN #17(00) 1382, Stockholm, Sweden, November 2000.

[4] "A^2IR – An asynchronous and adaptive HARQ scheme for HSDPA", Lucent Technologies, TSGR1#18(01) 0080, Boston, USA, January 2001.