New HARQ Scheme Based on Decoding of Tail-Biting Convolutional Codes in IEEE 802.16e

Hung-Ta Pai, *Member, IEEE*, Yunghsiang S. Han, *Fellow, IEEE*, and Yu-Jung Chu

***Abstract*—Traditionally, a hybrid automatic repeat request (HARQ) is executed at the physical (PHY) and medium access control (MAC) layers. A cyclic redundancy check (CRC) is usually performed at the MAC layer to decide whether a packet must be retransmitted. This design causes two problems—long latency and inefficient retransmission—when a transmission error appears in a large packet. In this paper, we propose a new HARQ scheme to solve these problems based not only on CRC but on the decoding of tail-biting convolutional codes (TBCCs) at the PHY layer as well. First, the TBCC codeword that was received at the decoder is cyclically shifted according to the reliability of the received code bits. The Viterbi algorithm (VA), starting from all states, is then applied. When every survival path in the VA is close enough to its corresponding abandoned path, retransmission of the codeword is invoked. Because retransmission is restricted to a codeword and determined at the PHY layer, short latency and efficient retransmission are achieved. Simulation results show that the proposed scheme reduces the number of retransmitted codewords by up to 43% compared with the traditional HARQ in IEEE 802.16e orthogonal frequency-division multiple access (OFDMA) under a Rayleigh faded channel when the 1024-point fast Fourier transform (FFT) and 64-state quadrature amplitude modulation (64-QAM) are employed.**

***Index Terms*—Automatic repeat request (ARQ), decoding, error-correcting codes (ECCs), hybrid automatic repeat request (HARQ), IEEE 802.16, Long Term Evolution (LTE), orthogonal frequency-division multiple access (OFDMA), reliability, tail-biting convolution codes, Worldwide Interoperability for Microwave Access (WiMAX).**

I. INTRODUCTION

FORWARD error correction (FEC) and automatic repeat request (ARQ)[1] are two error-control strategies that have been widely used in digital communications [1]. The former approach is implemented at the physical (PHY) layer, and the approach latter is implemented at the medium access control (MAC) layer. In a cross-layer design, these approaches can be combined to establish a hybrid automatic repeat request (HARQ) mechanism. Several advanced communication standards have adopted HARQ to build a reliable transmission link, including high-speed downlink/uplink packet access (HSDPA/HSUPA) [2], IEEE 802.16e orthogonal frequency-division multiple access (OFDMA) [3], and Third-Generation Partnership Project (3GPP) Long Term Evolution (LTE) [4].

There are several approaches for designing a HARQ scheme. In IEEE 802.16e OFDMA, a packet with cyclic redundancy check (CRC) is divided into several blocks, and each block is coded using tail-biting convolutional codes (TBCCs) at the PHY layer of the transmitter.[2] The receiver decodes every coded block at the PHY layer and assembles several blocks into a packet at the MAC layer. The ARQ mechanism decides whether the packet must be retransmitted according to the result of the CRC on the packet. If the CRC fails, the packet is incorrect and must be retransmitted from the transmitter. Chase combining [5] or incremental redundancy [6], [7] is then employed. When the receiver receives the retransmitted packet, it merges the retransmitted packet with the packet that was received in the previous run block by block and decodes the merged block at the PHY layer. Because the packet that was received in the previous run is well utilized, HARQ outperforms ARQ, particularly under poor communication conditions. However, retransmission is impossible before the CRC result is available. Thus, the latency of the packet transmission in this mechanism is long. Moreover, when only a few blocks of the received packet are incorrect, the retransmission of the entire packet wastes too much transmission bandwidth. In this paper, we propose a novel HARQ scheme based not only on the CRC but on the decoding of TBCCs as well, to shorten the latency in IEEE 802.16e, where TBCCs have been used to encode the data frame.

TBCC is one of the error-correcting codes (ECCs) used for FEC. Because TBCC can resolve the code rate loss in convolutional codes (CC) with zero tail, it has been adopted for FEC in IEEE 802.16e OFDMA and 3GPP LTE. Moreover, TBCCs are not only used in control packages but in data packages in IEEE 802.16e OFDMA as well. Several optimal and suboptimal algorithms have been proposed to decode TBCCs [8]–[12]. In [11], Wang *et al.* proposed to cyclicly extend received codewords by some fixed number of received bits and then decoded the resulting codewords by the Viterbi algorithm (VA). The wrap-around Viterbi algorithm (WAVA) [12] decoded received codewords by VA when they were cyclicly extended several

Manuscript received July 2, 2010; revised October 15, 2010 and January 1, 2011; accepted January 24, 2011. Date of publication February 4, 2011; date of current version March 21, 2011. This work was supported by in part the National Science Council (NSC) of the Republic of China, Taiwan, under Contract NSC 97-2221-E-305-001-MY2 and Contract NSC 96-2628-E-305-001-MY3. The review of this paper was coordinated by Dr. C. Ibars.

H.-T. Pai is with the Graduate Institute of Communications Engineering, National Taipei University, Taipei 237, Taiwan (e-mail: htpai@mail.ntpu.edu.tw).

Y. S. Han is with the Department of Electrical Engineering, National Taiwan University of Science and Technology, Taipei 106, Taiwan (e-mail: yshan@mail.ntust.edu.tw).

Y.-J. Chu is with Kcodes Corporation, Taipei 114, Taiwan (e-mail: s79469101@webmail.ntpu.edu.tw).

Digital Object Identifier 10.1109/TVT.2011.2111470

---

[1] ARQ is also known as automatic retransmission query.

[2] If a low-density parity check (LDPC) code is implemented in the PHY layer, it can perform error detection as traditional CRC does. Hence, determining packet (frame) validation is not necessarily implemented in the MAC layer. The proposed scheme is useful only when TBCCs are employed.

times of original lengths. Its performance is close to the optimal algorithm, with twice the lengths of the received codewords.

The final state of a TBCC encoder must be identical to its starting state, and thus, its codeword paths are circular [13]. Consequently, it is not necessary to start decoding from the first bit of its codeword, because every state could be the initial state. The starting bit can be determined according to the reliability of the code bit [14] such that a codeword can correctly be decoded with a higher probability using VA. This effort makes every survival path in VA as far away from its corresponding abandoned path as possible. Conversely, if every survival path is very close to its corresponding abandoned path, there is a high probability that the codeword will incorrectly be decoded [15]. This characteristic can be utilized to determine when the transmitter needs to resend the codeword. That is, the HARQ can be codeword (block) based at the PHY layer without reducing the error-correcting capability. No operation at the MAC layer is required when the decision to retransmit has been made in the PHY layer. Therefore, the latency in the original HARQ scheme can drastically be reduced. The key contributions of this paper are summarized as follows.

- Improve the TBCC decoding algorithm proposed in [11] by finding a good starting bit such that only 4.2% of the extra maximum complexity of WAVA is needed for similar error performance.
- Generalize the scheme proposed in [15] by relaxing the comparison between the survival and the abandoned paths in every path merge, and apply it to TBCCs.
- Propose a hybrid method for error detection such that the extra bandwidth for the CRC-based HARQ is reduced up to 43%.

The rest of this paper is organized as follows. Section II briefly addresses the related work. Section III presents the proposed scheme. A performance evaluation of the proposed scheme is given in Section IV. Conclusions are finally drawn, and future work is outlined in Section V.

II. PREVIOUS WORK

A. TBCC Decoding

Let $\mathcal{C}$ be an $(n, 1, m)$ TBCC of L information bits, where n is the number of output bits per information bit, and m is the memory order. Hence, the trellis of $\mathcal{C}$ is of $L+1$ levels (from the 0th to the Lth levels) and has 2^m states at each level. The corresponding tail-biting paths for the codewords of $\mathcal{C}$ should be constrained on the same initial and final states. Define the received vector as $\boldsymbol{r} \triangleq (r_0, r_1, \ldots, r_{nL-1})$.

Wang *et al.* [11] proposed applying VA to an extended received vector as

$$\boldsymbol{r}_{ext} = (r_0, r_1, \ldots, r_{nL-1}, r_0, r_1, \ldots, r_{\alpha nL-1})$$

where $\alpha > 0$. The length of $\boldsymbol{r}_{ext}$ is equal to $(1+\alpha)nL$, and the trellis becomes $(1+\alpha)L+1$ levels, from levels 0 to $(1+\alpha)L$. VA is applied to $\boldsymbol{r}_{ext}$ by equally likely initializing the metrics of all states at the zeroth level (usually equal to zero). When the final level is reached, the path with the best metric is traced back. Because the length of the traced codeword path is $(1+\alpha)L$, only the middle part of the codeword path, with length L, is decoded as the estimated codeword. Obviously, the algorithm can only find a suboptimal solution, and its performance is strongly related to α. A larger α implies a higher complexity and a lower word error rate (WER) at a fixed signal-to-noise ratio (SNR).

Because the TBCC codeword path is circular, a decoding algorithm can start from any level of the received codeword. Handlery *et al.* [14] proposed a reliability-based method to locate the starting level. Denote the binary codeword of $\mathcal{C}$ as $\boldsymbol{v} \triangleq (v_0, v_1, \ldots, v_{nL-1}) \in \{0,1\}^{nL}$. Define the bit reliability of code bit v_j as $|\phi_j|$, where

$$\phi_j \triangleq \ln \frac{p(r_j | v_j = 0)}{p(r_j | v_j = 1)}. \tag{1}$$

The starting level can be calculated by

$$\ell_{opt} = \arg\max_{\ell} \sum_{j=n\ell}^{n\ell+nw-1} |\phi_j| \tag{2}$$

where w is the window size. Obviously, this method can work with any TBCC decoding algorithm. It has been applied to different algorithms [10], [14] to the lower WER and/or computational complexity.

B. ECC Decoding With Repeat Request

ECCs can be used not only for FEC but also for error detection such that a retransmission request can be sent after the received vector is decoded. Yamamoto and Itoh modified VA for decoding zero-tail convolutional codes (ZTCCs) to determine whether the receiver should make a request for retransmission from the transmitter [15]. A node at level ℓ and state s is denoted as (ℓ, s), and a branch is represented by $(\ell-1, s') \to (\ell, s)$, where s' is a state at level $(\ell-1)$ such that the branch exists. For $(n, 1, m)$ CCs, including TBCCs and ZTCCs, exactly two paths merge at node (ℓ, s). Only the path with the best metric survives and is used for the remaining process, whereas the other path is discarded. Each node is marked with a label to indicate its uncertainty. The uncertainty of a node is determined by the metric difference between the survival path and the other going into this node. When all of the nodes at the same level are uncertain, the decoder at the receiver determines that the received codeword is unreliable and asks the transmitter for retransmission.

Then, Harvey *et al.* [16] improved the Yamamoto–Itoh (YI) method using packet-combining techniques. A packet is encoded as a single ZTCC codeword. If the YI method determines that the received packet (codeword) is unreliable, the transmitter must resend the same packet. The newly received codeword and the unreliably received codeword are combined. The combined packet, instead of the retransmitted packet alone, is then decoded. Because the ZTCC codeword is usually long, it is very possible that only a part of it is unreliable. An algorithm for detecting unreliable codeword segments was proposed by Freudenberger *et al.* [17]. Moreover, the YI method was applied to a trellis-coded modulation (TCM) system. A TCM-HARQ system has been presented and analyzed in [18]. However, this system does not perform as well as a system that uses CRC codes for error detection [19], [20].

Because CC is designed for FEC, not for error detection, it is not surprising that CC has worse error detection capability than CRC. For example, a CRC with 32 redundant bits (CRC-32) incorrectly detects an error with a very low probability, around $2^{-32} < 3 \times 10^{-10}$. Thus, it is impossible to replace CRC with CC for error detection. On the other hand, the error-detecting capability of CC using the YI method should be well utilized. In this paper, we proposed a new HARQ scheme that employs the capability to reduce the retransmission latency and bandwidth.

III. NEW HYBRID AUTOMATIC REPEAT REQUEST SCHEME

A. Analysis of Retransmission Bandwidth

Because TBCCs have no code rate loss, the length of a TBCC codeword can be short. When only a 1-b error in a codeword is detected in a codeword base, the codeword must be retransmitted. The retransmission bandwidth for only one codeword is needed. On the other hand, a packet usually consists of several TBCC codewords. When one of the codewords has only a 1-b error, the entire packet must be retransmitted if only CRC in the packet base is employed for error detection. Hence, the bandwidth for the retransmission is much higher than the bandwidth needed for the retransmission of one codeword.

Let k be the number of TBCC codewords that comprise a packet. N_w is the number of transmitted codewords. Assume that N_e codewords among these codewords are in error after decoding in the PHY layer. Thus, the number of packets is N_w/k, and the WER can be calculated by

$$\text{WER} = \frac{N_e}{N_w}.$$

Assuming that Chase combining is applied at the receiver, only one retransmission is required in almost all cases. Therefore, assuming that all codeword errors are independent of each other, if only CRC is used for error detection, the number of retransmitted codewords can be approximated by

$$\begin{aligned} N_r &\approx \left(1 - (1 - \text{WER})^k\right) \cdot \frac{N_w}{k} \cdot k \\ &= \left(1 - \left(1 - \frac{N_e}{N_w}\right)^k\right) N_w. \end{aligned}$$

If the YI method is employed for error detection in TBCC decoding, the number of retransmitted codewords can be approximated by

$$N_r' \approx \left(1 - \left(1 - \frac{\tilde{N}_e}{N_w}\right)^k\right) N_w + \tilde{N}_r$$

where $\tilde{N}_r$ is the number of retransmitted codewords based on the result of the YI method, and $\tilde{N}_e$ is the number of received codewords that is detected to be correct by the YI method but incorrect by the CRC. That is, $\tilde{N}_e$ erroneous codewords are missed by the YI method. When $N_e, \tilde{N}_e \ll N_w$, we have

$$\begin{aligned} N_r &\approx \left(1 - \left(1 - \frac{N_e}{N_w} k\right)\right) N_w = N_e k \\ N_r' &\approx \tilde{N}_e k + \tilde{N}_r. \end{aligned}$$

Define $\triangle N_e = N_e - \tilde{N}_e$ as the number of erroneous codewords received that are detected by the YI method. Therefore, if

$$\frac{\triangle N_e}{\tilde{N}_r} > \frac{1}{k}$$

then

$$N_r > N_r'.$$

That is, if the correct detection rate of the YI method is greater than $1/k$, the bandwidth for the retransmission can be reduced by jointly using the YI method and the CRC. This criterion can be used to decide when the YI method should be invoked during a decoding procedure. A detailed discussion will be presented in Section IV.

B. Reliability-Based Decoding of TBCC

Define the hard-decision sequence $\boldsymbol{y} = (y_0, y_1, \ldots, y_{nL-1})$ that corresponds to the received vector $\boldsymbol{r} = (r_0, r_1, \ldots, r_{nL-1})$ as

$$y_j \triangleq \begin{cases} 1, & \text{if } \phi_j < 0 \\ 0, & \text{otherwise} \end{cases}$$

where ϕ_j is defined in (1). Then, it can be found [21] that the ML decoding output $\boldsymbol{u} = (u_0, u_1, \ldots, u_{nL-1})$ for the received vector $\boldsymbol{r}$ satisfies

$$\sum_{j=0}^{nL-1} (u_j \oplus y_j)|\phi_j| \le \sum_{j=0}^{nL-1} (v_j \oplus y_j)|\phi_j| \tag{3}$$

for all $\boldsymbol{v} \in \mathcal{C}$, where "$\oplus$" is the exclusive-OR operation.

Because VA for a ZTCC starts its decoding from the all-zero state (0, 0), branches of the trellis of the ZTCC are not merged until the mth level. On the other hand, VA for a TBCC starts from all states with the same zero initial metric, i.e., $(0, s)$, $s = 0, 1, \ldots, 2^m - 1$, and branches of the trellis of the TBCC are merged right from the first level. That is, one of two branches, i.e., $(\ell - 1, s') \rightarrow (\ell, s)$ and $(\ell - 1, s'') \rightarrow (\ell, s)$, is selected as a survival path at node (ℓ, s) according to

$$M\left(P_{0,\ell}^{(s)}\right) = \min\left\{ M\left(P_{0,\ell-1}^{(s')}\right) + \sum_{j=2\ell}^{2\ell+1} \left(v_j' \oplus y_j\right)|\phi_j|, \; M\left(P_{0,\ell-1}^{(s'')}\right) + \sum_{j=2\ell}^{2\ell+1} \left(v_j'' \oplus y_j\right)|\phi_j| \right\} \tag{4}$$

where $P_{\ell_1,\ell_2}^{(s_2)}$ is the survival path from level ℓ_1 to node (ℓ_2, s_2), $M(\cdot)$ is the accumulated metric of a path, and $[v_{2\ell}' \; v_{2\ell+1}']$, $[v_{2\ell}'' \; v_{2\ell+1}'']$ are the code bits that correspond to the two

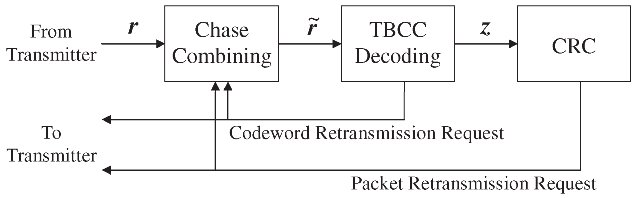

Fig. 1. Block diagram of the proposed scheme.

branches, respectively. This early merging of the branches may cause a high error probability. It is desirable to reduce the error probability. Based on (3), we observe that a larger $|\phi_j|$ can more explicitly differentiate two different bits u_j and v_j. This condition implies that the window size w in (2) can be set to m such that $(u_{n\ell_{opt}}, u_{n\ell_{opt}+1}, \ldots, u_{n\ell_{opt}+nm-1})$ can more easily be differentiated from $(v_{n\ell_{opt}}, v_{n\ell_{opt}+1}, \ldots, v_{n\ell_{opt}+nm-1})$. As a result, the received vector $\boldsymbol{r}$ is circularly left shifted by $n\ell_{opt}$ positions to get $\boldsymbol{r}' = (r_{n\ell_{opt}}, r_{n\ell_{opt}+1}, \ldots, r_{nL-1}, r_0, r_1, \ldots, r_{n\ell_{opt}-1})$, and the error probability caused by the early merging can be reduced by this setting.

Although the probability of correct decoding using VA is increased by choosing a better starting position for the received vector, it is still far away from the ML decoding performance [9]. The extension method that was proposed by Wang *et al.* [11] is adopted to increase the correct probability. That is, $\boldsymbol{r}'$ is extended to $\boldsymbol{r}'_{ext} = (r'_0, r'_1, \ldots, r'_{nL-1}, r'_0, r'_1, \ldots, r'_{\alpha nL-1})$, where $0 < \alpha \leq 1$, during decoding. VA is applied to $\boldsymbol{r}'_{ext}$ by starting from all states with the initial metric of zero. When the $((1+\alpha)L+1)$th level, i.e., the final level, is reached, 2^m survival paths are available. The final part, rather than the middle part [11], of the survival path, i.e., from levels (αL) to $((1+\alpha)L)$, is selected as the TBCC codeword path. Therefore, the survival path $P_{\alpha L,(1+\alpha)L}^{(s)}$ is a tail-biting path if the starting node of the path is $(\alpha L, s)$. That is, the starting and end nodes of a tail-biting path have the same state. The survival tail-biting path with the best metric is traced back, and then, the decoded bits are the ML solution with high probability. However, the ML solution is not guaranteed to be correct. In a harsh environment, the error probability of the ML solution is high, and retransmission of the codeword is necessary. The YI method can be applied to determine whether the receiver must ask the transmitter to again send the codeword.

C. Algorithm

We propose a new HARQ scheme, as illustrated in Fig. 1. The retransmission request can be sent by either the TBCC decoding at the PHY layer or the CRC at the MAC layer. When the retransmitted codeword (or packet) is received, it is combined with the previously received vector (or packet) $\boldsymbol{r}^{(p)}$ by Chase combining, i.e.,

$$\tilde{\boldsymbol{r}} = \frac{\boldsymbol{r}+\boldsymbol{r}^{(p)}}{2} = \left(\frac{r_0+r_0^{(p)}}{2}, \frac{r_1+r_1^{(p)}}{2}, \ldots, \frac{r_{nL-1}+r_{nL-1}^{(p)}}{2}\right)$$

where $\boldsymbol{r}^{(p)} \triangleq (r_0^{(p)}, r_1^{(p)}, \ldots, r_{nL-1}^{(p)})$. Equivalently, no Chase combining is performed in a new transmission, i.e., $\tilde{\boldsymbol{r}} = \boldsymbol{r}$. Notably, incremental redundancy, instead of Chase combining, may be applied.

Based on (4), denote the metrics of two paths that merge at node (ℓ, s) as

$$M'_{\ell,s} \triangleq M\left(P_{0,\ell-1}^{(s')}\right) + \sum_{j=2\ell}^{2\ell+1} \left(v'_j \oplus y_j\right)|\phi_j|$$

$$M''_{\ell,s} \triangleq M\left(P_{0,\ell-1}^{(s'')}\right) + \sum_{j=2\ell}^{2\ell+1} \left(v''_j \oplus y_j\right)|\phi_j|$$

respectively. Without loss of generality, assume that $M'_{\ell,s} < M''_{\ell,s}$. That is, the survival path $P_{0,\ell}^{(s)}$ is $P_{0,\ell-1}^{(s')}$ plus $(\ell-1, s') \rightarrow (\ell, s)$. Define T as a positive threshold. The TBCC decoding is executed as follows.

1) *Step 1*. Upon the reception of $\tilde{\boldsymbol{r}} = (\tilde{r}_0, \tilde{r}_1, \ldots, \tilde{r}_{nL-1})$ from the block of the Chase combining in Fig. 1, determine ℓ_{opt} according to (2), with $w = m$.
2) *Step 2*. Circularly left shift $\boldsymbol{r}$ by $n\ell_{opt}$ positions to get $\tilde{\boldsymbol{r}}' = (\tilde{r}_{n\ell_{opt}}, \tilde{r}_{n\ell_{opt}+1}, \ldots, \tilde{r}_{nL-1}, \tilde{r}_0, \tilde{r}_1, \ldots, \tilde{r}_{n\ell_{opt}-1})$, and extend $\tilde{\boldsymbol{r}}'$ to $\tilde{\boldsymbol{r}}'_{ext} = (\tilde{r}'_0, \tilde{r}'_1, \ldots, \tilde{r}'_{nL-1}, \tilde{r}'_0, \tilde{r}'_1, \ldots, \tilde{r}'_{\alpha nL-1})$, where $0 < \alpha \leq 1$.
3) *Step 3*. Decode $\tilde{\boldsymbol{r}}'_{ext}$ by starting from all states with the initial metric of zero.
4) *Step 4*. Set every node at level 0 with a label φ and $\ell = 1$.
5) *Step 5*. If $\ell > (1+\alpha)L$, go to *Step 8*. Otherwise, calculate the metric for each node, and process the trellis at level ℓ according to (4).
6) *Step 6*. For each node at level ℓ, if the difference between $M''_{(\ell,s)}$ and $M'_{(\ell,s)}$, i.e., $M''_{(\ell,s)} - M'_{(\ell,s)}$, is larger than T_ℓ and node $(\ell-1, s')$ has a label φ, then node (ℓ, s) has a label φ. Otherwise, node (ℓ, s) has a label χ.
7) *Step 7*. If all of the nodes at level ℓ have label χ, stop the decoding procedure, and ask the transmitter for retransmission. Otherwise, increase ℓ by one, and go to *Step 5*.
8) *Step 8*. If there is at least one tail-biting path among $P_{\alpha L,(1+\alpha)L}^{(s)}$, $s = 0, 1, \ldots, 2^m - 1$, the tail-biting path with the smallest metric is traced back. Otherwise, the path with the smallest metric is traced back. The information vector corresponding to the traced path is retained.
9) *Step 9*. Output information vector $\boldsymbol{z}$ by circularly right shifting the retained information vector in *Step 8* by ℓ_{opt} b.

Steps 1–3, *8*, and *9* are different from the original schemes [11], [14] in two ways. First, w in (2) is set to be m to reduce the error probability caused by early branch merging. Second, the survival path from levels $(1+\alpha)L$ to αL, rather than the middle part of the survival path, is traced back. The YI method is generalized to obtain *Steps 3–7*. The difference between two branches is not compared with the threshold until level $\alpha L + 1$. That is, only the traced-back part of the survival path is checked to determine whether the same part of any other path is close to it. Finally, because the input of the TBCC decoding is from the output of the demodulation, the value of the threshold is dependent on the modulation/demodulation methods and the transmission channels.

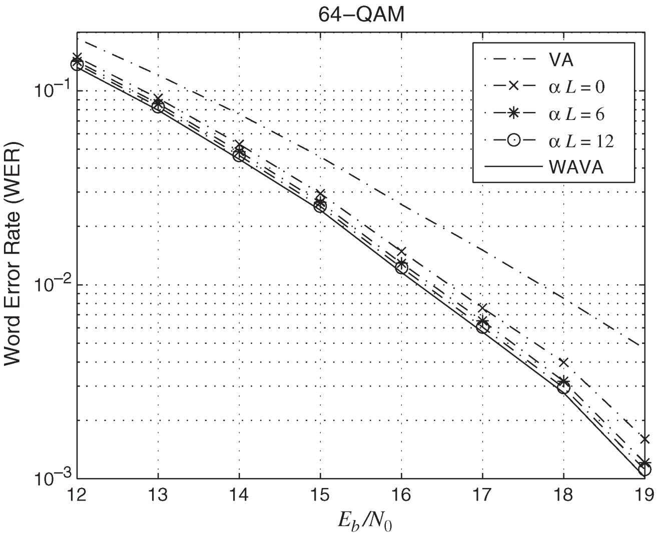

Fig. 2. Comparison among the Viterbi algorithm (VA), the wrap-around Viterbi algorithm (WAVA), and the proposed reliability-based TBCC decoding method using different αL for the (2, 1, 6) TBCC when the modulation scheme is 64-QAM.

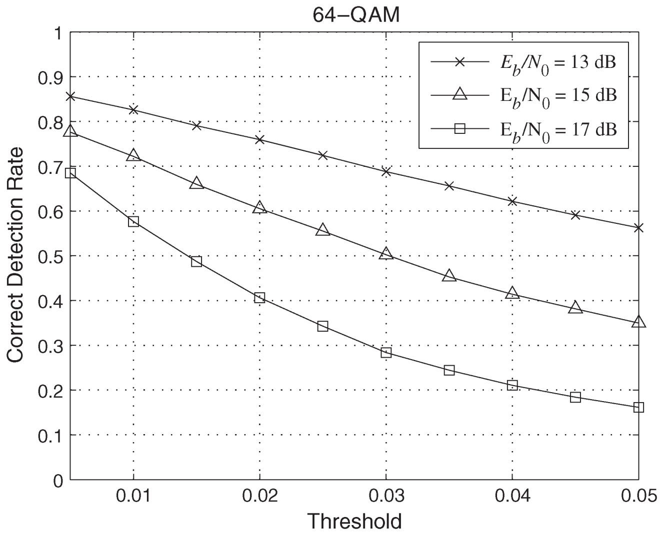

Fig. 3. Relation between the threshold and the correct detection rate of the proposed scheme in the cases of $E_b/N_0 = 13, 15,$ and 17 dB when the modulation scheme is 64-QAM.

IV. PERFORMANCE EVALUATION

We investigate the performance of the proposed scheme by simulations using IEEE 802.16e OFDMA with 1024 fast Fourier transform (FFT) points over the six-ray typical urban channel model with additive white Gaussian noise (AWGN) from COST 207 [22]. The (2, 1, 6) TBCC of 288 information bits is employed, and the generator polynomial of the code is (133, 171). The code rate is 1/2. The soft-decision output of demodulation is calculated using a simplified log-likelihood ratio (LLR) [23]. One protocol data unit (PDU) with CRC-32 includes five codewords, i.e., $k = 5$. The generating polynomial of CRC-32 is

$$x^{32}+x^{26}+x^{23}+x^{22}+x^{16}+x^{12}+x^{11}+x^{10}+x^{8} + x^{7}+x^{5}+x^{4}+x^{2}+x+1.$$

The number of collected decoding errors is at least 100. At most, one retransmission request for a codeword is allowed.

In the first set of simulations, 64-state quadrature amplitude modulation (64-QAM) is adopted because of its high modulation efficiency. All of the thresholds T_ℓ, $\ell = 1, 2, \ldots, (1+\alpha)L$ are set to be zero. That is, the YI method is not employed. The performance of VA, WAVA [12], and the proposed reliability-based TBCC decoding method using different αL are compared. WAVA may apply VA at most twice. Because $w = m = 6$ in (2), the 12 consecutive bits with the highest reliability in the codeword are utilized twice in the cases of $\alpha L = 6, 12$. Therefore, they perform very similar to WAVA, as shown in Fig. 2. In addition, the proposed scheme with $\alpha L = 0$ outperforms VA over 1 dB at 0.01 WER. This gain comes from choosing the starting level in (2), i.e., the reliability-based decoding. In the following simulations, αL is set to 12. That is, only $1 + 12/288 = 1.042$ runs of VAs are needed.

Next, we check whether the correct detection rate of the proposed scheme at the PHY layer can be greater than $1/k$, as discussed in Section III-A. Obviously, the error detection rate decreases as threshold T and the SNR increase, because the number of unnecessary retransmissions increases. That is, the proposed scheme frequently asks for retransmission for a correctly received codeword in the case of a large T and high SNR. When $E_b/N_0 = 13$ dB or 15 dB, the correct detection rate is greater than $1/5 = 0.2$ when T has values of 0.005–0.05, as illustrated in Fig. 3. On the other hand, when T has values of only 0.045–0.05 and $E_b/N_0 = 17$, the correct detection rate is lower than 0.2. Consequently, the correct detection rate is higher than $1/k$ in most cases, and then, the proposed scheme can outperform the traditional HARQ, i.e., using only CRC for error detection.

Fig. 4 shows a WER comparison between the proposed scheme $(T > 0)$ and the traditional HARQ $(T = 0)$. For fair comparison, the SNR is normalized as follows:

$$\text{Normalized} \quad E_b/N_0 = E_b/N_0 + 10\log_{10}(1 + \tilde{N}_r/N_w)$$

where $\tilde{N}_r$ is the number of retransmitted codewords based on the retransmission request of the PHY layer, and N_w is the number of transmitted codewords, as defined in Section III-A. As shown in Fig. 4, the proposed scheme significantly outperforms the traditional HARQ on WER for all T simulated. For example, when $T = 0.015$, the proposed scheme has a 2.3-dB coding gain over the traditional HARQ at WER $= 10^{-3}$. When $T = 0.035$, the coding gain is increased to 4 dB.

A larger threshold results in better WER performance. However, it also requires more bandwidth. Thus, a performance comparison between the proposed scheme and the traditional HARQ based on the reduced rate of retransmitted codewords is also illustrated in Fig. 5. The reduction rate of the retransmitted codewords at a fixed SNR is calculated by

$$1 - \frac{n'_r}{n_r}$$

where n_r and n'_r are the numbers of retransmitted codewords per codeword using the traditional HARQ and the proposed

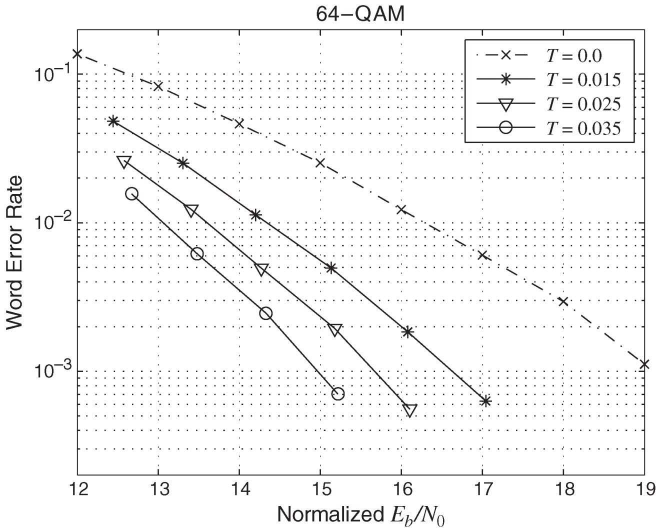

Fig. 4. WER comparison between the proposed scheme and the traditional HARQ using only CRC for error detection when the modulation scheme is 64-QAM.

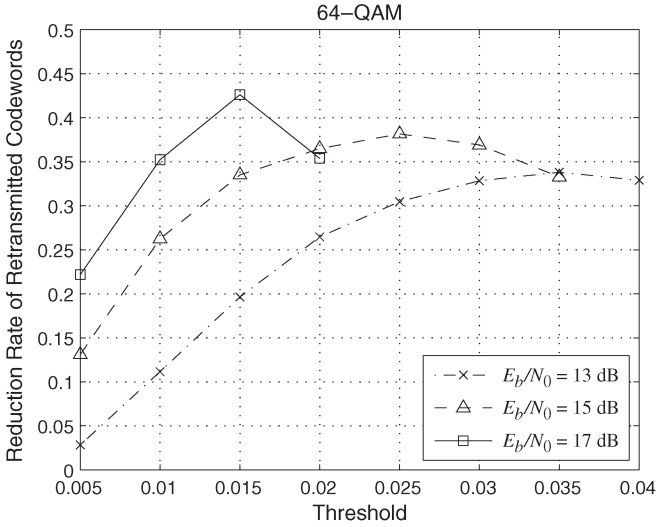

Fig. 5. Performance comparison between the proposed scheme and the traditional HARQ using only CRC for error detection when the modulation scheme is 64-QAM.

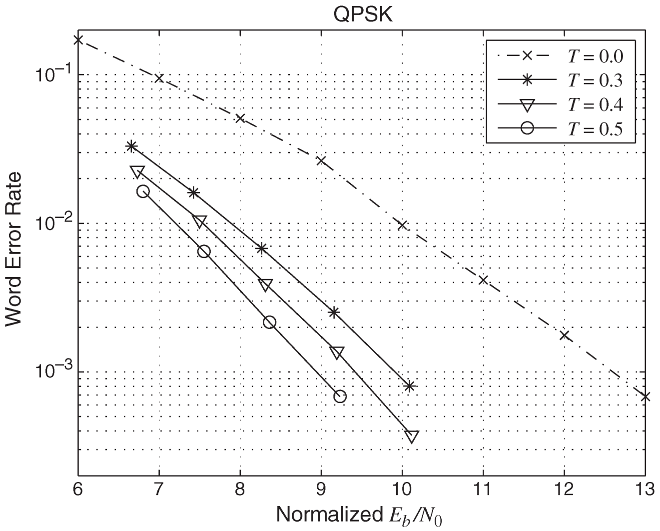

Fig. 6. WER comparison between the proposed scheme and the traditional HARQ using only CRC for error detection when the modulation scheme is QPSK.

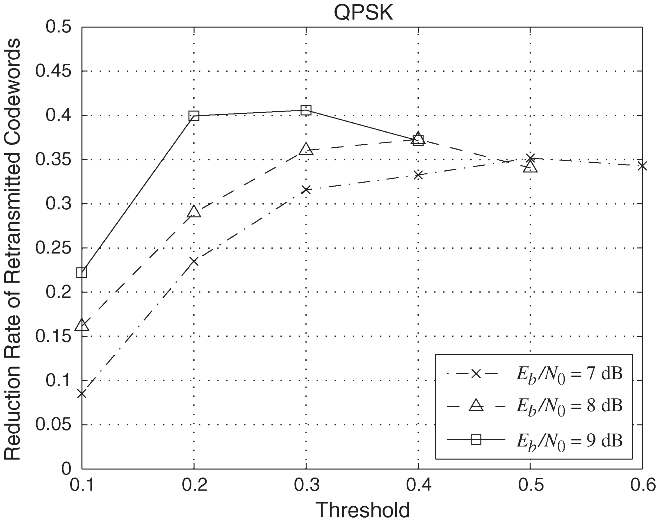

Fig. 7. Performance comparison between the proposed scheme and the traditional HARQ using only CRC for error detection when the modulation scheme is QPSK.

scheme, respectively. The optimal threshold T for each SNR is employed in the simulations. Fig. 5 shows that the number of retransmitted codewords can be reduced by up to 43% when $E_b/N_0 = 17$ dB and $T = 0.015$. At other operational SNRs, e.g., $E_b/N_0 = 13, 15$, the reduction rates are also more than 30%. Hence, the proposed scheme not only outperforms the traditional scheme in the WER performance but in the bandwidth efficiency as well. Note that the latency of the HARQ includes the time for the codeword retransmission and the decoding of the retransmitted codeword. That is, without considering the time for the packet processing in the MAC layer, the reduction in latency is proportional to the reduction in the retransmitted codewords.

In the second set of simulations, 64-QAM is replaced with quadrature phase-shift keying (QPSK) for a high-noise environment. Figs. 6 and 7 show that similar results are obtained. Hence, the proposed scheme also works very well in low-efficiency modulation.

V. CONCLUSION

This paper has proposed a new HARQ scheme that includes the Chase combining/incremental redundancy, reliability-based decoding of TBCCs, modified YI method, and CRC. Most retransmission requests are issued based on the decoding of TBCCs at the PHY layer so that short retransmission latency is obtained. Moreover, because the codeword, rather than the PDU, is retransmitted according the PHY-layer request, the retransmission bandwidth can be reduced. The proposed scheme outperforms the traditional HARQ, with a 43% reduction in the retransmission bandwidth and better WER performance under IEEE 802.16e OFDMA.

REFERENCES

[1] S. Lin and D. J. Costello, *Error Control Coding*. Englewood Cliffs, NJ: Prentice-Hall, 2004.

[2] H. Holma and A. Toskala, *HSDPA/HSUPA for UMTS: High-Speed Radio Access for Mobile Communications*. New York: Wiley, 2006.

[3] *IEEE Standard for Local and Metropolitan Area Networks Part 16—Air Interface for Fixed and Mobile Broadband Wireless Access Systems Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands and Corrigendum 1*, IEEE Std. 802.16e-2005, 2006.

[4] *3rd Generation Partnership Project; Tech. Spec. Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Multiplexing and Channel Coding (Rel. 8)*, 3GPP Std. TS 36.212 V8.6.0, 2009.

[5] D. Chase, "Code combining—A maximum-likelihood decoding approach for combining an arbitrary number of noisy packets," *IEEE Trans. Commun.*, vol. COM-33, no. 5, pp. 385–393, May 1985.

[6] D. Mandelbaum, "An adaptive-feedback coding scheme using incremental redundancy," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 3, pp. 388–389, May 1974.

[7] J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC codes) and their applications," *IEEE Trans. Commun.*, vol. 36, no. 4, pp. 389–400, Apr. 1988.

[8] P. Shankar, P. N. A. Kumar, K. Sasidharan, B. S. Rajan, and A. S. Madhu, *Efficient Convergent Maximum-Likelihood Decoding on Tail Biting*, Jan. 2006. [Online]. Available: http://arxiv.org/abs/cs.IT/0601023

[9] H.-T. Pai, Y. S. Han, T.-Y. Wu, P.-N. Chen, and S.-L. Shieh, "Low-complexity ML decoding for convolutional tail-biting codes," *IEEE Commun. Lett.*, vol. 12, no. 12, pp. 883–885, Dec. 2008.

[10] Y. S. Han, T.-Y. Wu, H.-T. Pai, P.-N. Chen, and S.-L. Shieh, "Priority-first search decoding for convolutional tail-biting codes," in *Proc. ISITA*, 2008, pp. 1–6.

[11] Y. P. E. Wang and R. Ramesh, "To bite or not to bite: A study of tail bits versus tail biting," in *Proc. PIMRC*, 1996, vol. 2, pp. 317–321.

[12] R. Y. Shao, S. Lin, and M. P. C. Fossorier, "Two decoding algorithms for tail-biting codes," *IEEE Trans. Commun.*, vol. 51, no. 10, pp. 1658–1665, Oct. 2003.

[13] R. V. Cox and C. E. W. Sundberg, "An efficient adaptive circular Viterbi algorithm for decoding generalized tail-biting convolutional codes," *IEEE Trans. Veh. Technol.*, vol. 43, no. 11, pp. 57–68, Feb. 1994.

[14] M. Handlery, R. Johannesson, and V. V. Zyablov, "Boosting the error performance of suboptimal tail-biting decoders," *IEEE Trans. Commun.*, vol. 51, no. 9, pp. 1485–1491, Sep. 2003.

[15] H. Yamamoto and K. Itoh, "Viterbi decoding algorithm for convolutional codes with repeat request," *IEEE Trans. Inf. Theory*, vol. IT-26, no. 5, pp. 540–547, Sep. 1980.

[16] B. A. Harvey and S. B. Wicker, "Packet combining systems based on the Viterbi decoder," *IEEE Trans. Commun.*, vol. 42, no. 2–4, pp. 1544–1557, Feb.–Apr. 1994.

[17] J. Freudenberger and B. Stender, "An algorithm for detecting unreliable code sequence segments and its applications," *IEEE Trans. Commun.*, vol. 52, no. 11, pp. 1833–1839, Nov. 2004.

[18] L. K. Rasmussen and S. B. Wicker, "The performance of type-I trellis-coded hybrid-ARQ protocols over AWGN and slowly fading channels," *IEEE Trans. Inf. Theory*, vol. 40, no. 2, pp. 418–428, Mar. 1994.

[19] L. K. Rasmussen and S. B. Wicker, "Trellis-coded type-I hybrid-ARQ protocols based on CRC error-detecting codes," *IEEE Trans. Commun.*, vol. 43, no. 10, pp. 2569–2575, Oct. 1995.

[20] L. K. Rasmussen and S. B. Wicker, "Comparison of two retransmission request mechanisms for trellis-coded modulation," *Electron. Lett.*, vol. 29, no. 25, pp. 2162–2163, Dec. 1993.

[21] J. Snyders and Y. Be'ery, "Maximum-likelihood soft decoding of binary block codes and decoders for the Golay codes," *IEEE Trans. Inf. Theory*, vol. 35, no. 5, pp. 963–975, Sep. 1989.

[22] G. L. Stüber, *Principles of Mobile Communication*. Boston, MA: Kluwer, 2001.

[23] F. Tosato and P. Bisaglia, "Simplified soft-output demapper for binary interleaved COFDM with application to HIPERLAN/2," in *Proc. ICC*, 2002, vol. 2, pp. 664–668.

Hung-Ta Pai (S'91–M'00) was born in Taichung, Taiwan. He received the B.Sc. degree in electrical engineering from the National Tsing Hua University, Hsinchu, Taiwan, in 1992 and the M.Sc. and Ph.D. degrees in electrical engineering from the University of Texas at Austin, in 1996 and 1999, respectively.

From 1992 to 1994, he served as an Army Officer. From 1999 to 2001, he was a Senior Design Engineer with Silicon Integrated Systems Corporation, Hsinchu. From 2001 to 2002, he was an Assistant Professor with the Department of Electrical Engineering, Tatung University, Taipei, Taiwan. From 2002 to 2004, he was an Assistant Professor with the Department of Electrical Engineering, National Taiwan University of Science and Technology, Taipei. Since August 2004, he has been with the National Taipei University, Taipei, where he is currently an Associate Professor with the Graduate Institute of Communications Engineering. His research interests include signal processing and communication systems.

Yunghsiang S. Han (F'11) was born in Taipei, Taiwan, on April 24, 1962. He received the B.Sc. and M.Sc. degrees in electrical engineering from the National Tsing Hua University, Hsinchu, Taiwan, in 1984 and 1986, respectively, and the Ph.D. degree from Syracuse University, Syracuse, NY, in 1993.

From 1986 to 1988, he was a Lecturer with Ming-Hsin Engineering College, Hsinchu. From 1989 to 1992, he was a Teaching Assistant with the School of Computer and Information Science, Syracuse University, where he was a Research Associate from 1992 to 1993. From 1993 to 1997, he was an Associate Professor with the Department of Electronic Engineering, Hua Fan College of Humanities and Technology, Taipei Hsien, Taiwan. From 1997 to 2004, he was with the Department of Computer Science and Information Engineering, National Chi Nan University, Nantou, Taiwan, where he was promoted to Professor in 1998. He was a Visiting Scholar with the Department of Electrical Engineering, University of Hawaii, Manoa, Honolulu, from June to October 2001, a Syracuse University Prototypical Research in Information Assurance Visiting Research Scholar with the Department of Electrical Engineering and Computer Science and CASE Center, Syracuse University, from September 2002 to January 2004, and a Visiting Scholar with the Department of Electrical and Computer Engineering, University of Texas at Austin, from August 2008 to June 2009. From August 2004 to July 2010, he was with the Graduate Institute of Communication Engineering, National Taipei University, Taipei. Since August 2010, he has been with the Department of Electrical Engineering, National Taiwan University of Science and Technology, Taipei. His research interests include error-control coding, wireless networks, and security.

Dr. Han received the 1994 Syracuse University Doctoral Prize.

Yu-Jung Chu was born in Taiwan on November 17, 1982. She received the B.S. degree in electrical engineering from the National University of Kaohsiung, Kaohsiung, Taiwan, in 2005 and the M.S. degree in communications engineering from the National Taipei University, Taipei, Taiwan, in 2008.

In 2009, she was an Instructor of electrical circuit analysis and signal analysis with the Air Force Institute of Technology, Kaohsiung. Since 2008, she has been a Software Development Engineer with Kcodes Corporation, Taipei, working on software implementation, and creating the interface for the common Internet file system (CIFS).