Seventh International Workshop on Optimal Codes and Related Topics
September 6-12, 2013, Albena, Bulgaria *pp. 119-124*

On the binary CRC codes used in the HARQ scheme of the LTE standard

Peter Kazakov peterkazakov@yahoo.com
Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, P.O.Box 323, 5000 Veliko Tarnovo, BULGARIA

Dedicated to the memory of Professor Stefan Dodunekov

Abstract. We investigate CRC codes generated by polynomials of degree $r = 24$ and minimum distance 4. Historically, standardized polynomials of degree r were chosen with a parity control check polynomial which ic product of $(x + 1)$ and a primitive polynomial of degree $r - 1$. We show that for the HARQ scheme of the LTE standard [5] this approach is far from optimal. We propose a method to select polynomials that perform better with respect to the function of probability of undetected error for specific codelengths. Moreover, we do not need exhaustive search to find better polynomials.

1 Introduction

Let C be a binary $[n_c, k_c, d = 4]$ CRC code generated by the polynomial $g(x)$. We recall that for each polynomial $g(x)$ there is a number n_c, such that $g(x)$ divides $x^{n_c}+1$ and $n_c = \min\{m|x^m \equiv 1 \pmod{g(x)}\}$. The number n_c is called order of the polynomial $g(x)$ and is denoted by $ord(g)$. So, length of code C is n_c and $\deg(g) = r = n_c - k_c$. Each codeword $c(x)$ can be represented as a product $a(x)g(x)$, where $\deg(a) < n_c - r$.

For a binary symmetric channel (BSC) the probability of undetected error can be expressed in the following way

$$P_{ud}(C, \varepsilon) = \sum_{i=1}^{n_c} A_i \varepsilon^i (1 - \varepsilon)^{n-i},$$

where ε is the channel error rate and $\{A_i\}_{i=0}^{n_c}$ is the distance distribution of the code C.

So, not only the minimum distance is important to characterize a certain CRC code with respect to probability of undetected error, but also the number of minimum weight codewords. This characteristic is especially important when ε is close to zero. Let us denote the number of minimum weight codewords by $A_{d,n_c-s}(g)$ for a shortened in s positions $[n_c - s, k_c - s, d = 4]$ CRC code. Consistently with the previous definition, we have $A_d(g) = A_{d,n_c}(g)$.

We will say that a CRC code of length n_c has an optimum value of $A_{d,n_c}(g)$, if it is the smallest possible one for this length.

A NP-complete method of finding a polynomial $g(x)$ that generates a code with minimum value of $A_{4,n_c-s}(g)$ is presented in [8] and used in [1] and [2]. In summary, for a given s, we need to calculate values of $A_{4,n_c-s}(g)$ for at most 2^{r-1} polynomials. This can be done by calculating generator matrix of size r and counting the weights of all possible combinations. Their number is 2^r and this will give us the dual distance distribution. We can get the value of $A_{4,n_c-s}(g)$ by applying the Mac-Williams transformation [6].

2 Classes of polynomials

Definition 1. *If two polynomials g(x) and f(x) can be factorized on an equal number k of irreducible polynomials* $g_1(x), \ldots, g_k(x)$ *and* $f_1(x), \ldots, f_k(x)$ *such that* $\deg(g_i(x)) = \deg(f_i(x))$ *for* $i = 1, \ldots, k$ *and* $ord(g_i(x)) = ord(f_i(x))$ *for* $i = 1, \ldots, k$ *we will say that they belong to one class.*

All polynomials from one class generate equivalent cyclic codes with the same $A_d(g)$.

This work is inspired by two (classes of) polynomials with optimal $A_{d,n_c-s}(g)$ for big intervals of codelengths for a fixed $r = 16$. Their optimal range is given in [7]. The first of these (classes of) polynomials is optimal for the interval $8002 \leq n_c - s \leq 19685$ and the second one for the interval $1286 \leq n_c - s \leq 8001$. Note that both polynomials have odd weight, i.e. $(x+1)$ does not divide them. These polynomials are:

$$\begin{aligned} g_1(x) &= x^{16} + x^{15} + x^{14} + x^{12} + x^{10} + x^3 + 1 \\ &= (x^4 + x^3 + x^2 + x + 1)(x^5 + x^4 + x^3 + x + 1)(x^7 + x^6 + 1), \\ g_2(x) &= x^{16} + x^{15} + x^{12} + x^9 + x^8 + x^6 + x^5 + x^4 + 1 \\ &= (x^3 + x^2 + 1)(x^6 + x^5 + x^2 + x + 1)(x^7 + x^6 + x^5 + x^3 + x^2 + x + 1) \end{aligned}$$

For them it holds $ord(g_1) = 3 * 31 * 127 = 19685$ and $ord(g_2) = 63 * 127 = 8001$. To give indication about the properties of these polynomials, we compare $A_d(g_i), i = 1, 2$, with $A_d(f)$ of a shortened in the corresponding number of positions CRC code generated by the polynomial f of degree 16 and order $2^{15} - 1$. In our case we use the CRC-16-CCITT standardized polynomial $f(x) = x^{16} + x^{12} + x^5 + 1$. We have in Table 1 comparison of the corresponding A_d values.

So, we see that these polynomials have a considerably better A_{4,n_c}. Since the case $r = 16$ is covered in [7] we are interested to see how such CRC codes behave for other $r = 24$. In particular, we compare these polynomials to the

Code length	Optimal polynomials	Standard polynomial
8001	$A_{4,8001}(g_2) = 3,550,443,750$	$A_{4,8001}(f) = 5,205,812,877$
19685	$A_{4,19685}(g_1) = 158,813,048,515$	$A_{4,19685}(f) = 190,889,834,302$

Table 1: Comparison of A_d of standardized polynomials and optimal ones of degree 16.

LTE standard. The LTE standard [5] defines two polynomials, namely:

$$g_A(x) = x^{24} + x^{23} + x^{18} + x^{17} + x^{14} + x^{11} + x^{10} + x^7 + x^6 + x^5 + x^4 + x^3 + x + 1$$

and

$$g_B(x) = x^{24} + x^{23} + x^6 + x^5 + x + 1.$$

Error correction schemes in LTE standard are described extensively in [3] and [4]. In brief, the standard uses blocks of maximum 6144 bits, 6120 information bits and 24 CRC bits generated by polynomial g_A. All transmitted information is protected additionally with 24 CRC bits generated by polynomial g_B. Both polynomials have order $8,388,607$.

3 Method of investigation

To determine the number of minimum weight codewords of an extended Hamming code we can refer to

Theorem 1. *([7, Theorem 9]) Let C be a binary $[n_c - r, k_c, 4]$ code generated by the polynomial $(x+1)g(x)$ of degree r and order $n_c = 2^{r-1} - 1$. Then the following equality holds:*

$$A_{4,n_c-s}(g) = \frac{(n_c-3)((n_c-4s)(n_c-1)+6s(s-1))}{24} - B, \tag{1}$$

where $B = \sum_{m=2}^{max(2,s-1)} \sum_{j=1}^{m-1} (s-m)(Q_{m,j}(g) - \sum_{l=1}^{j-1} Q_{m,j,l}(g))$,

$$Q_{m,j}(g) = \begin{cases} 1, & g(x) \mid x^m + x^j + 1, \\ 0, & otherwise \end{cases}$$

$$Q_{m,i,j}(g) = \begin{cases} 1, & g(x) \mid x^m + x^j + x^i + 1, \\ 0, & otherwise. \end{cases}$$

According to Theorem 1, the Hamming codes shortened in s positions generated by primitive polynomials of degree $r-1$ multiplied by $x+1$ have values of $A_{4,n_c-s}(g)$ in a close range because the most terms in B from (1) will be 0 or 1.

Unfortunately, we do not have such formulas for an arbitrary CRC code. In this work we use several techniques to easily obtain generator polynomials of CRC codes that perform significantly better, if not optimal, for a given $n_c - s$.

The classical method of calculation of $A_{4,n_c-s}(g)$ is to use the distance distribution of the dual code calculated with the standard Gray code method [9] and then to apply the MacWillams transformation [6]. The idea behind the Gray code is that each codeword is generated at a simple step based on the current state and the previous codeword. Therefore the complexity of this algorithm is $O(2^r)$, $r = ord(g)$. If we apply exhaustive search, we need to do this for all polynomials except reciprocal ones, so an additional factor of 2^{r-1} applies.

In our method, we still use Gray code to calculate dual distance distribution, however, we do this on significantly smaller number of polynomials. It can be summarized in four steps.

1. We group all polynomials in classes according to Definition 1. We exclude reciprocal polynomials, i.e. $g(x)$ and $x^{24}g(1/x)$, since they have the same order and distance distribution. So, we take only one of them.

2. For each order n_c and factorization, we select one polynomial h from every class and we calculate the minimum distance d of the corresponding CRC code. If $d = 3$, we skip this class of polynomials.

3. For each class represented by a polynomial h, we calculate the number of minimum weight codewords $A_{4,6120}(h)$ and select two groups of polynomial classes - one with a polynomial representative of order bigger than 6120 and one of order bigger than $100,000$.

4. For the first three classes with a minimum value of $A_{d=4,6120}$ (in order to compare with g_A) and the first three classes with a minimum value of $A_{d=4,100,000}$ and an order bigger than $100,000$ (in order to compare with g_B) and big codelengths we perform calculations on all their members. In that way we find the best polynomial from the corresponding class that generates a minimum $A_{d=4,6120}$.

The software computational modules are developed by the author and are available on request.

4 Results

In the tables below we give our results for the investigated degrees. We propose new polynomials that generate CRC codes with a much smaller number of minimum weight codewords than the polynomials used in the LTE standard. All polynomials are presented in hexadecimal notation, for example the polynomial $x^{24} + x^6 + x^5 + x^4 + x^3 + 1$ is denoted by 0x1000079 and $x^4 + x^3 + x + 1$ is denoted by 0x1B.

Polynomial notation	*order*	$A_{d,6120}$
0x1864CFB (standard,g_A)	$2^{23} - 1$	56,416,496
0x114855B	38227	24,989,800
0x17A481F	12291	25,013,640
0x14AC147	19065	25,463,304

Table 2: Comparison of the minimum weight codewords of standardized polynomials and new proposals with order lower than 100,000.

Polynomial notation	*order*	$A_{d,6120}$
0x1800063 (standard,g_B)	$2^{23} - 1$	68,018,112
0x103A977	114681	25,201,272
0x116C3EF	522753	25,850,512
0x140F133	278845	29,275,776

Table 3: Comparison of the minimum weight codewords of the standardized polynomials and the new proposals of order higher than 100,000.

We notice that all optimal polynomials have odd weight and they perform significantly better than the corresponding standard polynomial.

5 Conclusions

In this work we propose polynomials which perform significantly better than the standard LTE polynomials for the target codelenght. We group polynomials in classes and we select ones with $d = 4$ and minimum value of A_{4,n_c-s} for $n_c - s = 6120$ and $100,000$. The proposed polynomials perform much better compared to standardized LTE polynomials with respect to the function of undetected error probability for the most useful case when ε is close to zero. Although the whole algorithm is NP-complete, exhaustive calculation of 2^r polynomials is replaced

by one of limited number of investigated polynomials. We have shown that for large range of codelengths we can obtain easily polynomials that perform better with respect to the probability of undetected error.

References

[1] Ts. Baicheva, S. Dodunekov, P.Kazakov, On the cyclic redundancy-check codes with 8-bit redundancy, *Computer Communications*, 21, 1998, 1030-1033.

[2] Ts. Baicheva, S. Dodunekov, P.Kazakov, Undetected error probability performance of cyclic redundancy-check codes of 16-bit redundancy, *IEE Proceedings – Communications* 147, 2000, 253-256.

[3] F. Berkmann, C. Carbonelli, F. Dietrich, C. Drewes, W. Xu, On 3G LTE terminal implementation – standard, algorithms, complexities and challenges, Proc. IWCMC, 2008.

[4] Jung-Fu (Thomas) Cheng, H. Koorapaty, Error detection reliability of LTE CRC coding, Vehicular Technology Conference, 2008, VTC 2008-Fall.

[5] ETSI TS 136 212 V9.3.0 (2010-10) Technical Specification LTE.

[6] F. J. MacWilliams, A theorem on the distribution of weights in a systematic code, *Bell Syst. Tech. J.*, 42, 1963, 79-94.

[7] P. Kazakov, Fast calculation on the number of minimum weight words of CRC codes, *IEEE Trans. on Inform. Theory* 47, March 2001, 1190-1195.

[8] P. Kazakov, Application of Polynomials to CRC and Spherical Codes, PhD Thesis, TU Delft, The Netherlands, 2000.

[9] W. Lipski, Kombinatoryka dla programistow, Wydawnictwa Naukowo-Techniczne, Warszawa, 1982.