Fast Calculation Algorithm of the Undetected Errors Probability of CRC Codes

Ren-Der Lin and Wen-Shyen Chen
Department of Computer Science, National Chung-Hsing University
Taichung, Taiwan 402, ROC
lrd@cs.nchu.edu.tw , echen@cs.nchu.edu.tw

Abstract

The error detecting functions of linear block code can be realized via simple software or hardware. The error detection, which includes long-term theoretical research and many good properties, is often applied widely in digital communication and data storage. The weight distribution of linear block code and its dual code are important parameters of calculating the probability $\mathbb{P}_{ud}$ of undetected errors. Further, Cyclic Redundancy Check (CRC) codes and Bose, Chaudhuri and Hocquenghem (BCH) cyclic codes are subclasses of linear block codes. This paper proposes a fast calculation algorithm of weight distribution of the dual code which outperforms those of previous studies in time complexity, and the probability of undetected error of different CRC codes standards under various codeword lengths are also simulated efficiently.

Keywords: *linear block code*, *the dual code*, *weight distribution*, *probability of undetected error.*

1. Introduction

A_i and B_i are the number of codewords of weight i in (n, k, d) linear block code C and its dual code $C^{\perp}$, so $\{A_i\}_{i=0}^{n}$ and $\{B_i\}_{i=0}^{n}$ are the weight distribution of C and $C^{\perp}$, respectively.

$$A(z) = A_0 + A_1 z + \cdots + A_n z^n$$

$$B(z) = B_0 + B_1 z + \cdots + B_n z^n$$

The 2^k codewords are involved in the weight distribution $\{A_i\}_{i=0}^{n}$ of C, e.g. (n, k) = (1040, 1024), $2^k = 2^{1024} \fallingdotseq 1.8*10^{308}$. But only 2^m codewords are involved in the weight distribution $\{B_i\}_{i=0}^{n}$ of $C^{\perp}$, e.g. $m = n - k = 1040 - 1024 = 16$, $2^m = 2^{16} = 65536$. Obviously, $\{B_i\}_{i=0}^{n}$ is much easier to be calculated than $\{A_i\}_{i=0}^{n}$.

CRC codes are often (but not always) constructed of the polynomial $\overline{g}(x) = (1+x)\,p(x)$, where $p(x)$ is a primitive polynomial of degree r that divides $x^{2^r-1}-1$. An even weight code polynomial of C has $(1+x)$ as a factor.

To iSCSI, the error detection capability of a selected detection mechanism should be proper at least up to block lengths of 8k bytes (64kbits).**[1]**

To Binary Symmetric Channels (BSCs), bit error rate $\varepsilon \in [0, 0.5]$, the exact (non-bounding) evaluation of probability $\mathbb{P}_{ud}$ of undetected error of linear block code is defined as follows:

$$\mathbb{P}_{ud}(C,\varepsilon) = \sum_{i=1}^{n} A_i \varepsilon^i (1-\varepsilon)^{n-i}$$

$$= 2^{-m} \sum_{i=0}^{n} B_i (1-2\varepsilon)^i - (1-\varepsilon)^n$$

The performance evaluation of a linear block code depends on the probability of undetected errors. $\mathbb{P}_{ud}$ is the popular appraisal parameter to decide whether linear block code codes are good or proper. Since A_i and B_i both can obtain $\mathbb{P}_{ud}$, to calculate B_i is much easier than A_i. Further, the calculated complexity of B_i can be greatly reduced via the fast calculation algorithm of B_i. This paper will focus here. This paper also proposes a fast algorithm to improve the performance evaluation.

Proceedings of the 19th International Conference on Advanced Information Networking and Applications (AINA'05)
1550-445X/05 $20.00 © 2005 IEEE

2. Previous work

In recent years, this topic of fast calculation algorithms is mentioned by some methods which focus on the fast calculation algorithm of weight distribution $\{B_i\}_{i=0}^{n}$ of $C^{\perp}$ as follows:

2.1. The method [2] from Tohru Fujiwara, Tadao Kasami, Atsushi Kitai, and Shu Lin. The recursive method must be restricted on factor $(1+x)$ of generator polynomial $\overline{g}(x)$, i.e. $\overline{g}(x) = (1+x)\ \mathrm{p}(x)$, where $\mathrm{p}(x)$ must be primitive generator polynomial for a cyclic code.

2.2. The method **[3]** from Guy Castagnoli, Stefan Bräuer, and Martin Herrmann present extension of **[2]**. Although there is no restriction on factor $(1+x)$ of generator polynomial $g(x)$ by the method. But the class of code must be also cyclic codes, and the time complexity is about $O(2^{2m})$ pointed by **[5]**for codeword lengths limited between m and n_c, where n_c is the maximum codeword length at $d > 2$. However, the time complexity is higher than the previous one.

2.3. The method **[4]** from Chun, D.; Wolf, J.K. Special hardware of two identical LFSR for efficiently evaluating the weight distribution of the dual codes generated by $\overline{g}(x) = (1+x)\ p(x)$, where $p(x)$ is a primitive polynomial can be determined for cyclic code of any n up to 65535. The time complexity becomes $O(2^r)$.

2.4. The direct method. For linear combinations of the rows of parity check matrix H can generate the codewords of $C^{\perp}$, is then used to calculate the weight distribution. The direct method is as reference for our recursive method in next chapter.

3. Fast calculation algorithm of weight distribution of the dual code

This recursive calculation algorithm will compute codeword length n from that of codeword length n - 1. Because the length of paper cannot derive enough, mathematics is skipped in this final camera-ready paper.

This method bases on two key points shown below:

3.1. Re-design P_λ

G^{sys} is important parameter for this paper. Linear block code (both cyclic and non-cyclic) can be made systematic by matrix row operations that the time complexity will be variable. To non-cyclic codes, G^{sys} can be made generally by Gaussian elimination that this paper will not introduce.

This paper conducts a simply recursive algorithm for cyclic codes based on dependent relation.

3.2. Re-design $C^{\perp}$

This paper also conducts a quick recursive algorithm of the parity check matrix H based on dependent relation.

Proceedings of the 19th International Conference on Advanced Information Networking and Applications (AINA'05)
1550-445X/05 $20.00 © 2005 IEEE

Table 1 Comparison of some fast calculation algorithm of weight distribution of the dual code

	Class of codes	Restriction on the factorization of the generator polynomial	Restriction on the generator polynomial is a primitive polynomial	Maximum codeword length	Complexity per codeword
Our method	All linear codes	None	None	$> n_c$	$O\left(2^r\right)$
[2]	Only cyclic codes	V	V	32767	$O\left(2^r\right)$
[3]	Only cyclic codes	None	None	n_c	$O\left(2^{2r}\right)$
[4]	Only cyclic codes	V	V	65535	$O\left(2^r\right)$

4. The results of simulation

In **Fig. 1**, X axis denotes code lengths m+1 ~ 65536; Y axis denotes the best standards of CRC-16. We present the best standards which attain minimum values of $\mathbb{P}_{ud}$ for different values of n. Although the best standard of CRC-16 is the same under various bit error rate ε, it is different for the codeword lengths. In fact, these standards indicate whether they are suitable for current application.

In **Fig. 2 ~ Fig. 4**, X axis denotes code lengths m+1 ~ 65536; Y axis denotes $\mathbb{P}_{ud}$. Thus, it shows clearly that the $\mathbb{P}_{ud}$ performance is really different under various CRC-16 standards or various codeword lengths n. Also this presents the importance for selecting the proper standard.

5. Concluding remarks

In this paper, we propose the fast calculation algorithm of weight distribution $\{B_i\}_{i=0}^{n}$ of the dual code $C^{\perp}$. It efficiently in time complexity solves the original difficult computation. Compared with the previous algorithms, as these constants are variable, the complexity is similar to **[2]**. But our advantages that there is no restriction on factor $(1+x)$ of generator polynomial $g(x)$, this method is more efficient in time complexity than **[3]** and the direct method. And this method isn't also limited for cyclic codes; it can be applied for any linear block codes as **Table 1**. The proposed algorithm greatly reduces the time complexity. When applying these rules to quickly compare the performances of each CRC standard, the best standard can be selected effectively, and even new best generator polynomial can be discovered. These new standards are very expectable.

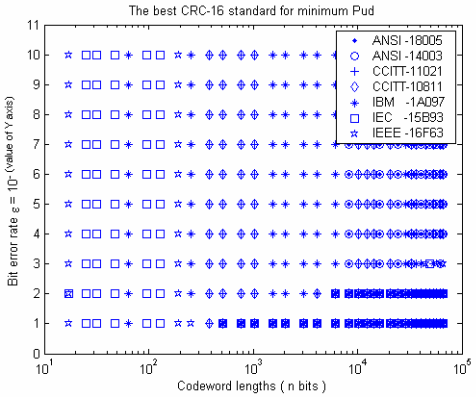

Fig. 1. The best CRC-16 standard under some codeword lengths for bit error rate $\varepsilon = 10^{-1}, 10^{-2}, \cdots, 10^{-10}$.

Proceedings of the 19th International Conference on Advanced Information Networking and Applications (AINA'05)
1550-445X/05 $20.00 © 2005 IEEE

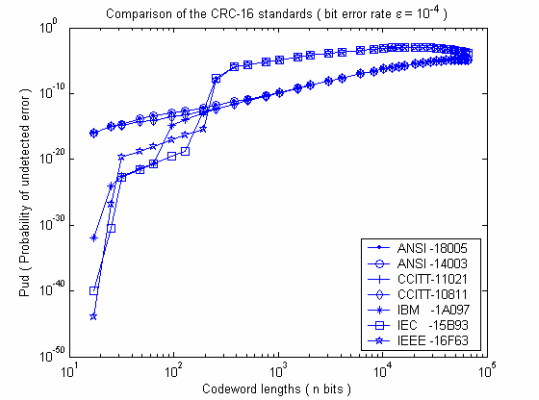

Fig. 2. The probability $\mathbb{P}_{ud}$ of undetected error of the CRC-16 standards for bit error rate $\varepsilon = 10^{-4}$.

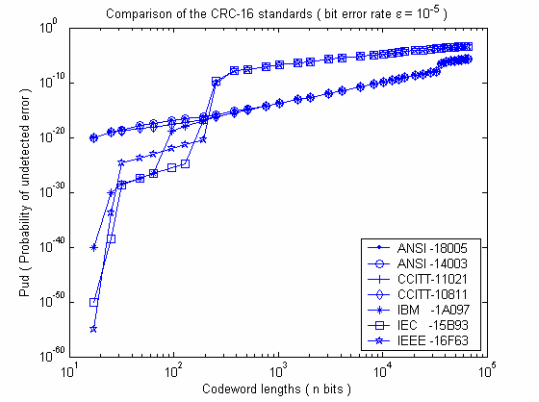

Fig. 3. The probability $\mathbb{P}_{ud}$ of undetected error of the CRC-16 standards for bit error rate $\varepsilon = 10^{-5}$.

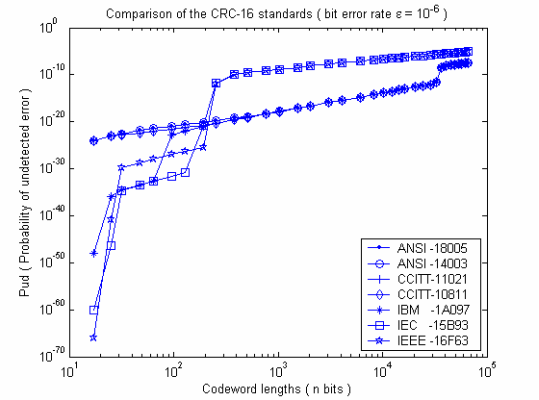

Fig. 4. The probability $\mathbb{P}_{ud}$ of undetected error of the CRC-16 standards for bit error rate $\varepsilon = 10^{-6}$.

6. Acknowledgements

The authors wish to thank Dr. Peter Kazakov and Professor Martin Röder for their useful discussions.

7. References

[1] Meth K.Z., and Satran J. , "Features of the iSCSI protocol", *IEEE Communications Magazine*, Vol. 41, Aug. 2003, pp. 72 - 75.

[2] Tohru Fujiwara, Tadao Kasami, Atsushi Kitai, and Shu Lin, "On the Undetected Error Probability for Shortened Hamming Codes", *IEEE Trans. on Communications* Vol. 33, June 1985, pp. 570-574.

[3] Guy Castagnoli, Stefan Bräuer, and Martin Herrmann, "Optimization of cyclic redundancy check codes with 24 and 32 parity bits", *IEEE Trans. on Communications* Vol. 41, June 1993, pp. 883-892.

[4] Chun, D., and Wolf, J.K., "Special hardware for computing the probability of undetected error for certain binary CRC codes and test results", *IEEE Trans. on Communications* Vol. 42, Oct. 1994, pp. 2769-2772.

[5] Peter Kazakov, "Fast calculation of the number of minimum-weight words of CRC codes", *IEEE Trans. on Information Theory* Vol. 47, March 2001, pp. 1190-1195.