International Symposium on Information Theory and its Applications, ISITA2004
Parma, Italy, October 10–13, 2004

A Novel Modification of Cyclic Redundancy Check for Message Length Detection

Shin-Lin Shieh[†‡], Po-Ning Chen[‡] and Yunghsiang S. Han[*]

[†] Computer & Comm. Research Labs, Industrial Technology Research Institute, Chutung Township, HsinChu County, Taiwan 31041, R.O.C.
E-mail:shinlin@micro.ee.nthu.edu.tw

[‡] Dept. of Communications Engineering, National Chiao-Tung University, Hsin Chu City, Taiwan 30056, R.O.C.

[*] Dept. of Computer Science and Information Engineering, National Chi Nan University, Nan Tou, Taiwan 545, R.O.C.

Abstract

Cyclic redundancy check (CRC) bits that are conventionally used for error detection have recently found a new application in UMTS WCDMA standard (specifically, "blind transport format detection") for message length detection of variable-length message communications. Co-worked with the inner convolutional code, it was demonstrated that the CRC bits can simultaneously detect the length of a message block, which was unknown to the receiver, without much degradation in its error detection capability. In order to improve the correct probability of length detection, NTT DoCoMo proposed to reverse the transmission order of CRC bits, and showed by simulations that the false-length probability can be flattened to a small constant value for all wrong lengthes. In this work, we provided an analytical proof to the flatten false-length probability of NTT DoCoMo reverse-CRC scheme. In addition, we proposed an alternative modification of the original CRC method by selectively flipping some of the CRC bits, and analytically proved that the false-length probability of our modification can be made exactly zero for every length offset (defined as the difference between the true length and the detected length) smaller than the number of CRC bits, if the inner convolutional code has corrected all channel errors. A necessary and sufficient condition under which a polynomial can serve as a flip polynomial for CRC bits with length detection enhancement is also established.

1. INTRODUCTION

It is quite common in many communication systems that the length of transmitted message blocks varies. Additional length information is thus necessary for the receiver to de-block the message. To achieve a reliable transmission, an error correcting code is often used to protect the message where a fixed number of CRC bits attached at the end for a possible retransmission if some uncorrected channel errors do occur (cf. Fig. 1(a)).

In some specific applications, the data rate is prohibitively low that the transmission of additional length information becomes an inefficient system burden. An example is the AMR 12.2 kbps mode of UMTS WCDMA, in which the transmission overhead for message length may be as large as 3 kbps, which is almost 25% of the 12.2 kbps data rate. In such case, detection of message length through the attached CRC bits with the help of inner convolutional code decoder becomes a potential system alternative. Notably, a length is accepted as a legal candidate only when a certain node on the trellis of the convolutional code gives the smallest metric among all nodes at the same level, and at the same time, the validity test of the CRC bits is passed (cf. blind transport format detection using CRC in pages 56–58 of [3] or the system diagram in Fig. 1(b)).

To facilitate the analytical study of length detection capability of CRC bits, we consider the false-length probability of CRC bits by assuming that all channel errors have been corrected by the inner convolutional code. It can be shown that directly applying the standard CRC method for message length detection will introduce high false-length probability at small length offset. In [2], NTT DoCoMo proposed a modification on the standard CRC, and empirically showed that the false-length probability can be brought down to a small constant of $2^{-\ell}$ for all wrong lengths, where ℓ is the number of attached CRC bits. In this work, we confirmed DoCoMo's simulation result by an analytical proof.

To further improved DoCoMo's reverse-CRC scheme, we introduce another modification of the standard CRC method by selectively flipping some of the

This work was supported in part by Chung-Shan Institute of Science & Technology (CSIST), Taiwan, ROC, under Grant XC93B95P.

CRC bits, and analytically proved that the false-length probability can be further reduced to zero for every message length offset smaller than the number of CRC bits. Since our modification of CRC bits, as well as DoCoMo's modification, is simply a one-to-one correspondence between the modified CRC bits and the original CRC bits, it can be shown by simulations that the error detection capability from the system view of Fig. 1(b) (in which CRC test is passed only when the convolutional decoder simultaneously reaches the desired state) remains almost intact. A necessary and sufficient condition under which a polynomial can serve as a flip polynomial for a specific CRC mechanism with length detection enhancement is also derived.

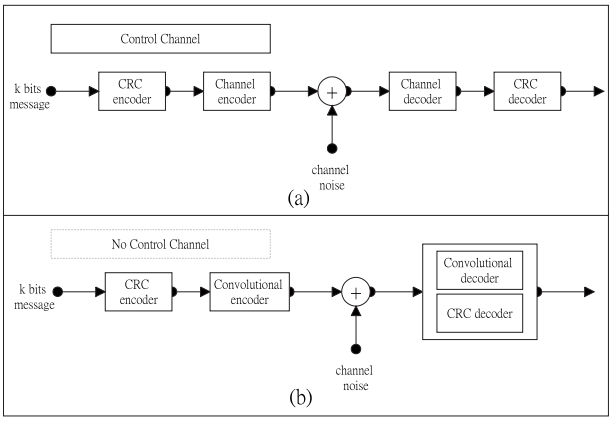

Figure 1: (a) Block diagram of a typical communication system, in which the conceptual "Control Channel" block carries the length information to the receiver. (b) System diagram of the WCDMA system in "blind transport format detection" mode, in which the convolutional decoder and CRC decoder should co-work to extract the length information that was not sent by the transmitter, as indicated by the dot-framed "No Control Channel".

2. Standard CRC Methods

A usual condition for the CRC generating polynomial $g_\ell(x)$ of order ℓ is that $\gcd(g_\ell(x), x^i) = 1$ for each $0 \le i \le \ell$. Some popular CRC generating polynomials that satisfy this condition are quoted (from [1][3]) as follows.

- CRC-8 : $g_8(x) = x^8 + x^7 + x^4 + x^3 + x + 1$
- CRC-12 : $g_{12}(x) = x^{12} + x^{11} + x^3 + x^2 + x + 1$
- CRC-16 : $g_{16}(x) = x^{16} + x^{12} + x^5 + 1$
- CRC-24 : $g_{24}(x) = x^{24} + x^{23} + x^6 + x^5 + x + 1$

As aforementioned, the CRC bits can also be applied to validate the message block length. For clarity, the encoding and decoding procedures are explicitly listed in the following.

Standard CRC: Fix a pre-specified CRC generating polynomial $g_\ell(x)$ of order ℓ.

- ***Encoder***
  1. For a given message block $M = [m_{k-1}, \ldots, m_0]$ of k bits, determine ℓ parity check bits $P = [p_{\ell-1}, \ldots, p_0]$ such that $g_\ell(x) \,|\, \big(x^\ell M(x) + P(x)\big)$, where $M(x) = m_{k-1}x^{k-1} + \cdots + m_0$ and $P(x) = p_{\ell-1}x^{\ell-1} + \cdots + p_0$.
  2. Attach these ℓ parity check bits at the end of the k message bits to form a coded block of $C = [m_{k-1}, \ldots, m_0, p_{\ell-1}, \ldots, p_0]$ for transmission.
- ***Decoder***
  1. After the reception of a message block $R = [r_{\hat{k}+\ell-1}, \ldots, r_0]$ of length $\hat{k}+\ell$, calculate the ℓ parity check bits $[\hat{p}_{\ell-1}, \ldots, \hat{p}_0]$ for the surmised message block $[r_{\hat{k}+\ell-1}, \ldots, r_\ell]$ (by following the same procedure as the encoder).
  2. If $[\hat{p}_{\ell-1}, \ldots, \hat{p}_0] = [r_{\ell-1}, \ldots, r_0]$,
     then the CRC test for message length $\hat{k}$ is passed (and a legitimate message block is found);
     else append the next input bit to R to form a new message block of length $\hat{k}+1+\ell$, and repeat Steps 1∼2 with a new $\hat{k} = \hat{k} + 1$.

Our first next lemma shows that without the help of the inner convolutional code, the standard CRC will introduce large false-length probability even if all message bits are corrected transmitted.

Lemma 1. *Under the assumptions of (i) uniformly distributed message, (ii) noise-free transmission and (iii) that the message length is no smaller than twice of the CRC size, the false-length probability $P_F(i)$ of the standard CRC is given by:*

$$P_F(i) = \begin{cases} 0, & \text{for } i = 0; \\ 2^{-i}, & \text{for } 1 \le i \le \ell - 1; \\ 2^{-\ell}, & \text{for } i \ge \ell, \end{cases}$$

where $i = k - \hat{k}$ is the message length offset.

Proof:

1. *Claim* : If $p_0 = 0$, then the CRC test will be passed with $i = 1$. On the contrary, if $p_0 = 1$, then the CRC test will be failed with $i = 1$.

The claim can be proved as follows. The received message with length offset $i = 1$ is equal to $R = [m_{k-1}, \ldots, m_0, p_{\ell-1}, \ldots, p_1]$. Now if $p_0 = 0$, the coded block for transmission equals:

$$\begin{aligned} C(x) &= m_{k-1}x^{m+\ell-1} + \cdots + p_1 x \\ &= x(m_{k-1}x^{m+\ell-2} + \cdots + p_1) \\ &= xR(x), \end{aligned}$$

Since $g_\ell(x)|C(x)$ and $\gcd(g_\ell(x), x) = 1$, we obtain $g_\ell(x)|R(x)$.

On the contrary, if $p_0 = 1$,

$$\begin{aligned} C(x) &= m_{k-1}x^{m+\ell-1} + \cdots + p_1 x + 1 \\ &= x(m_{k-1}x^{m+\ell-2} + \cdots + p_1) + 1 \\ &= xR(x) + 1, \end{aligned}$$

which, by $g_\ell(x)|C(x)$ and $\gcd(g_\ell(x), x) = 1$, immediately gives that $g_\ell \nmid R(x)$.

2. For uniformly distributed message, $p_0 = 0$ and $p_0 = 1$ are equally probable. Thus, $P_F(1) = 2^{-1}$.

3. It can be proved by following similar procedure as **1.** that only when $(p_0, p_1) = (0, 0)$ can $g_\ell(x)$ divide $m_{k-1}x^{k+\ell-3} + \cdots + m_0 x^{\ell-2} + p_{\ell-1}x^{\ell-3} + \cdots + p_2$. Since $(p_0, p_1) = (0, 0)$ with probability $1/4$, $P_F(2) = 2^{-2}$.

4. The proof for other values of i with $i \le \ell - 1$ can be completed by using the same reasoning as above; hence, details are omitted.

5. $P_F(i) = 2^{-\ell}$ for $i \ge \ell$, since the probability of uniformly distributed $[m_{i-1}, \ldots, m_{i-\ell}]$ equal to the parity bits generated due to the message bits of $[m_{k-1}, \ldots, m_i]$ is equal to $2^{-\ell}$. □

3. DoCoMo's Modification of CRC Methods

In order to reduce the false-length probability when the number of length offset is smaller than the CRC size, NTT DoCoMo proposed to reverse the order of the attached CRC bits.

DoCoMo's Modification: Fix a pre-specified CRC generating polynomial $g_\ell(x)$ of order ℓ.

- ***Encoder***

  1. For a given message block $M = [m_{k-1}, \ldots, m_0]$ of k bits, determine ℓ parity check bits $P = [p_{\ell-1}, \ldots, p_0]$ such that $g_\ell(x) \,\big|\, \big(x^\ell M(x) + P(x)\big)$, where $M(x) = m_{k-1}x^{k-1} + \cdots + m_0$ and $P(x) = p_{\ell-1}x^{\ell-1} + \cdots + p_0$.
  2. Attach these ℓ parity check bits *in reverse order* at the end of the k message bits to form a coded block of $C = [m_{k-1}, \ldots, m_0, p_0, \ldots, p_{\ell-1}]$ for transmission.

- ***Decoder***

  1. After the reception of a message block $R = [r_{\hat{k}+\ell-1}, \ldots, r_0]$ of length $\hat{k} + \ell$, calculate the ℓ parity check bits $[\hat{p}_{\ell-1}, \ldots, \hat{p}_0]$ for the surmised message block $[r_{\hat{k}+\ell-1}, \ldots, r_\ell]$ (by following the same procedure as step 1 of the encoder).
  2. If $[\hat{p}_{\ell-1}, \ldots, \hat{p}_0] = [r_0, \ldots, r_{\ell-1}]$,
     then the CRC test for message length $\hat{k}$ is passed (and a legitimate message block is found);
     else append the next input bit to R to form a new message block of length $\hat{k}+1+\ell$, and repeat Steps 1∼2 with a new $\hat{k} = \hat{k} + 1$.

Next lemma shows the false-length probability of DoCoMo's method.

Lemma 2. *Under the assumptions of (i) uniformly distributed message, (ii) noise-free transmission and (iii) that the message length is no smaller than 2 times of the CRC size, the false-length probability $P_{F,rev}(i)$ of the reverse CRC is given by:*

$$P_{F,rev}(i) = \begin{cases} 0, & \text{for } i = 0; \\ 2^{-\ell}, & \text{for } i > 0, \end{cases}$$

where $i = k - \hat{k}$ is the message length offset.

Proof:

1. For $i \ge \ell$ and $i = 0$, the result is clear. Therefore, we may assume $0 < i < \ell$.
2. Let the CRC-coded block be $[m_{k-1}, \cdots, m_0, p_{\ell-1}, \cdots, p_0]$ such that
   $$g_\ell(x) | m_{k-1}x^{k+\ell-1} + \cdots + m_0 x^\ell + p_{\ell-1}x^{\ell-1} + \cdots + p_0.$$
3. According to DoCoMo's proposal, the transmitter transmits $[m_{k-1}, \cdots, m_0, p_0, \cdots, p_{\ell-1}]$. In the decoder, if the length offset is i, $0 < i < \ell$, the received vector is
   $$R = [m_{k-1}, \cdots, m_i, m_{i-1}, \cdots, m_0, p_0, \cdots, p_{\ell-1-i}].$$
4. The CRC test will check if $g_\ell(x)$ divides $R_{rev}(x)$ where
   $$R_{rev} = [m_{k-1}, \cdots, m_i, p_{\ell-1-i}, \cdots, p_0, m_0, \cdots, m_{i-1}],$$

and

$$R_{rev}(x) = m_{k-1}x^{k+\ell-1-i} + \cdots + m_i x^{\ell} + p_{\ell-1-i}x^{\ell-1} + \cdots + p_0 x^i + m_0 x^{i-1} + \cdots + m_{i-1}.$$

5. Reverse CRC test will pass if, and only if,

$$\begin{aligned} & g_\ell(x) \mid R_{rev}(x), \\ \Leftrightarrow\ & g_\ell(x) \mid x^i R_{rev}(x), \\ \Leftrightarrow\ & g_\ell(x) \mid (m_{k-1}x^{k+\ell-1} + \cdots + m_i x^{\ell+i}) \\ & \quad + (p_{\ell-i-1}x^{\ell+i-1} + \cdots + p_0 x^{2i}) \\ & \quad + (m_0 x^{2i-1} + \cdots + m_{i-1}x^i), \\ \Leftrightarrow\ & g_\ell(x) \mid (m_{i-1}x^{\ell+i-1} + \cdots + p_0) \\ & \quad + (p_{\ell-i-1}x^{\ell+i-1} + \cdots + p_0 x^{2i}) \\ & \quad + (m_0 x^{2i-1} + \cdots + m_{i-1}x^i). \end{aligned}$$

Let $Q(x) = q_{\ell+i-1}x^{\ell+i-1} + \cdots + q_0 = (m_{i-1}x^{\ell+i-1} + \cdots + m_0 x^{\ell} + p_{\ell-1}x^{\ell-1} + \cdots + p_0) + (p_{\ell-i-1}x^{\ell+i-1} + \cdots + p_0 x^{2i}) + (m_0 x^{2i-1} + \cdots + m_{i-1}x^i)$.

6. Since degree of $Q(x)$ is no greater than $\ell + i - 1$, the number of choices of $Q(x)$ such that $g_\ell(x)|Q(x)$ is 2^i with $1 < i < \ell$.

7. Express the relation of $Q(x)$ and $C(x)$ in matrix form,

$$\begin{bmatrix} q_{\ell+i-1} \\ \vdots \\ q_\ell \\ q_{\ell-1} \\ \vdots \\ q_0 \end{bmatrix} = \mathbf{M} \begin{bmatrix} m_{i-1} \\ \vdots \\ m_0 \\ p_{\ell-1} \\ \vdots \\ p_0 \end{bmatrix}$$

where

$$\mathbf{M} = \mathbf{I}_{(\ell+i)} + \begin{bmatrix} \mathbf{0}_{(\ell-i)\times i} & \mathbf{0}_{(\ell-i)\times i} & \mathbf{I}_{(\ell-i)} \\ \mathbf{J}_i & \mathbf{0}_i & \mathbf{0}_{i\times(\ell-i)} \\ \mathbf{0}_i & \mathbf{0}_i & \mathbf{0}_{i\times(\ell-i)} \end{bmatrix},$$

$\mathbf{I}_\ell$ is an $\ell \times \ell$ identify matrix, and all components of $\mathbf{J}_\ell$ are zeros except those locate at ith row and $(\ell - i)$th column, which equal unity.

8. Since the matrix $\mathbf{M}$ is invertable, there exists a one-to-one correspondence between $[q_{\ell+i-1}, \cdots, q_0]$ and $[m_{i-1}, \cdots, m_0, p_{\ell-1}, \cdots, p_0]$. Therefore, there are also 2^i choices of $[m_{i-1}, \cdots, m_0, p_{\ell-1}, \cdots, p_0]$ such that $g_\ell(x)|R_{rev}(x)$.

9. Under the assumption of uniform distributed message and that the message length is no smaller than twice of the CRC size, the probability that $g_\ell(x)|R_{rev}(x)$ is

$$2^i \times 2^{-(\ell+i)} = 2^{-\ell} \text{ for } 1 < i < \ell.$$

□

4. Our Proposed Modification of Standard CRC

The main idea of the proposed modification is to selectively flip some CRC bits. Specifically, for a given ℓ-bit CRC, we construct a flip polynomial of order $\ell-1$, denoted by $f_\ell(x) = t_{\ell-1}x^{\ell-1} + \cdots + t_0$, where $t_i \in \{0, 1\}$ for $0 \le i \le \ell - 1$. Then, the jth parity bits p_j is "flipped" when $t_j = 1$, and "unflipped", otherwise. For clarity, our CRC encoding and decoding rules are provided below.

The proposed CRC Modification: Fix a prespecified CRC generating polynomial $g_\ell(x)$ (that satisfies $\gcd(g_\ell(x), x^i) = 1$ for each $0 \le i \le \ell$) and its corresponding flip polynomial $f_\ell(x)$.

- ***Encoder***

  1. For a given message block $M = [m_{k-1}, \ldots, m_0]$ of k bits, determine ℓ parity check bits $P = [p_{\ell-1}, \ldots, p_0]$ such that $g_\ell(x) \,|\, \big(x^\ell M(x) + P(x)\big)$, where $M(x) = m_{k-1}x^{k-1} + \cdots + m_0$ and $P(x) = p_{\ell-1}x^{\ell-1} + \cdots + p_0$.
  2. Flip the ℓ parity check bits according to the flip polynomial $f_\ell(x)$. The resultant parity check vector is $[\bar{p}_{\ell-1}, \ldots, \bar{p}_0] = [p_{\ell-1} + t_{\ell-1}, \ldots, p_0 + t_0]$, where "+" represents modulo-2 addition operator.
  3. Attach the flipped ℓ parity check bits at the end of the k message bits to form a coded block of $C = [m_{k-1}, \ldots, m_0, \bar{p}_{\ell-1}, \ldots, \bar{p}_0]$ for convolutional encoding.

- ***Decoder***

  1. After the reception of a convolutional decoded message block $R = [r_{\hat{k}+\ell-1}, \ldots, r_0]$ of length $\hat{k} + \ell$, calculate the ℓ parity check bits $[\hat{p}_{\ell-1}, \ldots, \hat{p}_0]$ for the surmised message block $[r_{\hat{k}+\ell-1}, \ldots, r_\ell]$ (by following the same procedure as step 1 of the encoder).
  2. If $[\hat{p}_{\ell-1} + t_{\ell-1}, \ldots, \hat{p}_0 + t_0] = [r_{\ell-1}, \ldots, r_0]$ (and a certain condition for the convolutional decoding is verified), then the CRC

test for message length $\hat{k}$ is passed (and a legitimate message block is found and recorded for a possible candidate);

3. Append the next input bit to R to form a new message block of length $\hat{k}+1+\ell$, and repeat Steps 1∼2 with a new $\hat{k}=\hat{k}+1$ until the end of the frame is reached.

For a specific ℓ-bit CRC generating polynomial $g_\ell(x)$, the necessary and sufficient condition under which $f_\ell(x)$ can make the false-length probability $P_{F,\text{Flip}}(i)=0$ for every length offset $0\le i\le \ell-1$ is established below.

Lemma 3. *Under uniformly distributed message and error-free transmission,*

$$P_{F,\text{Flip}}(i)=\begin{cases}0 & \text{, for } 0\le i\le \ell-1;\\ 2^{-\ell} & \text{, for } i\ge \ell,\end{cases}$$

if, and only if,

$$\deg\left(\text{Remainder of }\left\{\frac{(1+x^i)f_\ell(x)}{g_\ell(x)}\right\}\right)\ge i \quad (1)$$

for $1\le i\le \ell-1$.

Proof: Let $D(x)=m_{k-1}x^{k+\ell+1}+\cdots+m_0x^\ell+p_{\ell-1}x^{\ell-1}+\cdots+p_0$. Then the polynomial corresponding to the coded block C is equal to $C(x)=D(x)+f_\ell(x)$.

At the receiver end, let $R^{[i]}(x)$ be the polynomial corresponding to a wrong message size of $k-i$, i.e.,

$$R^{[i]}(x)=\frac{D(x)+f_\ell(x)+\sum_{j=0}^{i-1}(p_j+t_j)x^j}{x^i}.$$

Hence, the CRC test for message length $k-i$ will be passed if, and only if, $g_\ell(x)\,\big|\big(R^{[i]}(x)+f_\ell(x)\big)$. Since $\gcd(g_\ell(x),x^i)=1$ for $0\le i\le \ell$, the condition of $g_\ell(x)\,\big|\big(R^{[i]}(x)+f_\ell(x)\big)$ is equivalent to

$$g_\ell(x)\,\Bigg|\left(D(x)+f_\ell(x)+\sum_{j=0}^{i-1}(p_j+t_j)x^j+x^if_\ell(x)\right). \quad (2)$$

We then observe that (2) is valid if, and only if,

$$g_\ell(x)\,\Bigg|\left\{\sum_{j=0}^{i-1}(p_j+t_j)x^j+(1+x^i)f_\ell(x)\right\}, \quad (3)$$

because $g_\ell(x)|D(x)$. Accordingly, the validity of (1) implies that (3) cannot hold for every $1\le i\le \ell-1$, because the order of $\sum_{j=0}^{i-1}(p_j+t_j)x^j$ is at most $i-1$. This gives that $P_{F,\text{Flip}}(i)=0$ for $1\le i\le \ell-1$.

On the contrary, if (1) fails for some $1\le i\le \ell-1$, then the remainder polynomial $\lambda(x)$ of $g_\ell(x)$ dividing $\sum_{j=0}^{i-1}t_jx^j+(1+x^i)f_\ell(x)$ is of the order, at most, $i-1$. Hence, there must exist some $(p_0,\ldots,p_{i-1})$ such that $\sum_{j=1}^{i-1}p_jx^j$ plus $\lambda(x)$ being zero, which, by the assumption of uniformly distributed message, implies $P_{F,\text{Flip}}(i)>0$. □

With the availability of the necessary and sufficient condition, we can exhaustively search all the legitimate flip polynomials by computers. For 8-bit CRC protection with $g_8(x)=x^8+x^7+x^4+x^3+x+1$, the number of qualified flip polynomials is 74.

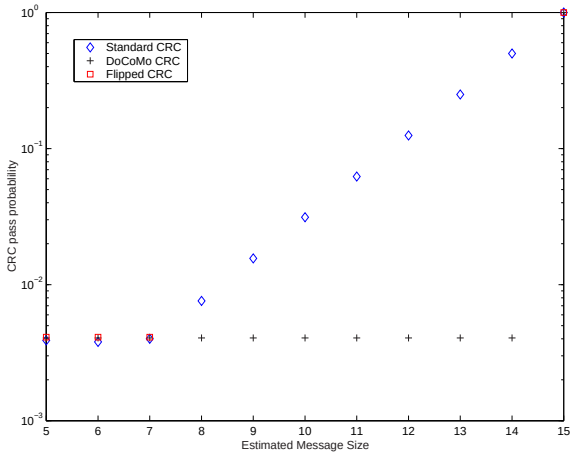

Figure 2: Simulations for CRC pass probabilities for each message size with true message size $k=15$ and CRC generating polynomial $g_8(x)=x^8+x^7+x^4+x^3+x+1$. The flip polynomial used here is $f_8(x)=x^7+1$.

5. Conclusions and Future Work

In this work, we proposed a new modification of the CRC method by selectively flipping some of the CRC bits according to a pre-specified flip polynomial. We also analytically proved that if all channel errors are corrected by the inner convolutional code, the false-length probability of our new modification can be made zero for every length offset smaller than the number of CRC bits. The derived necessary and sufficient condition makes it possible to exhaust all the candidate flip polynomials by computers. An interesting future work would be to examine and compare the system performances of these candidates when convolutional code and channel errors are introduced.

References

[1] S. B. Wicker, *Error Control Systems for Digital Communication and Storage*. Prentice Hall, 1995.

[2] NTT DoCoMo, *TSGR1#5(99)689* TSG-RAN Working Group1 meeting #5, June 1999.

[3] 3rd Generation Partnership Project, "Multiplexing and channel coding(FDD)," 3GPP Tech. Spec., TS 25.212 V3.9.0, March 2002.