A Performability Aspect of Tandem CRCs’ Probability of Undetectable Burst Errors

MENG-LAI YIN
Electrical and Computer Engineering, California State Polytechnic University, Pomona, USA, and Raytheon Network Centric Systems, Specialty Engineering, Fullerton, USA

(Received on August 17, 2009, revised on January 1, 2010)

Abstract – In this article, the probability of undetectable burst errors for network systems applying the tandem Cyclic Redundancy Check (CRC) coding scheme is analyzed under a performability framework.

Keywords: Cyclic redundancy check code, burst errors.

1. Introduction

For error-detection Cyclic Redundancy Check (CRC) coding scheme, a performance metric being concerned is the probability of undetectable burst errors. This attribute is of particular interest in safety-critical systems, such as the satellite navigation systems for aircrafts landing, where the safety requirement of misleading information is extremely small, e.g., in the order of $10^{-9}$. For these systems, corrupted information is only one factor considered in the overall safety concern; hence the requirement of probability of undetectable burst errors is even smaller. Yet the probability of undetectable errors for a 24-bit CRC is in the order of $10^{-8}$ [3]. Tandem CRCs was first proposed in [2] to extend the guaranteed detectable burst error length, which is then used to support the safety-critical systems for the stringent requirement mentioned above.

Conceptually, when two CRC codes are used in tandem, a second CRC code is appended to the end of the message with the first CRC code. Denote a message with length $d$ in a polynomial form as $M(x)$, and the two CRC general polynomials as $g(x)$ and $h(x)$, with lengths $m$ and $n$ respectively. The coding process is first dividing $M(x)X^m$ by $g(x)$, which gives the first CRC code, followed by dividing $M(x)X^m+R_g(x)$ by $h(x)$, which yields the second CRC code. The decoding mechanism first checks the received message with $h(x)$. If the message passed the first check, a second decoding scheme applies which checks the highest $(d+m)$ bits by $g(x)$ [2]. An undetectable error pattern must pass both checks.

To show that the requirements are met for the aforementioned applications, it is inevitable to analyze the probability of undetectable errors. In [4], an analysis method based on the positions and the lengths of the burst errors was conducted; however, this method is tedious and error-prone. In this article, a different approach is taken which takes advantages of the fact that CRC implements a hash function and the probability of undetectable burst errors can be analyzed under a performability framework.
2. A Performability Approach

Performability combines two aspects, *i.e.*, performance and dependability [1], where performance usually means how well a system can work, and dependability gives the idea of how well the system would support a particular performance. In this paper, the performance considered is the probability of undetectable burst errors; for dependability, it is addressed by the occurrences of the error patterns.

Theoretically, performability is often described by a stochastic process \( Y=\{Y_t|t \in T\} \) [1]. In this paper, \( T \) is discrete and \( t \) is the time a message is transmitted. The random variable \( Y \), represents the performance of the coding scheme at time \( t \), and is defined by domain \( S \) and co-domain \( R \), *e.g.*, \( Y: S \to R \). Specifically, \( S \) is the set of all possible error patterns, while \( R \) is the “reward” which indicates whether an error pattern is detectable or not. If the generator polynomials are known, it is determinable (through a hash function) whether an error pattern is detectable.

A commonly used performability metric is the *expected reward rate* [1]. Let’s denote the reward for an error pattern \( s \) as \( r_s \), where \( r_s \) is either 0 (the error pattern \( s \) is undetectable) or 1 (otherwise). If the probability that error pattern \( s \) would occur is denoted as \( p_s \), then the expected reward rate is \( \sum_{s \in S} p_s \times r_s \). This metric is the probability of detectable burst errors, or *one* minus the probability of undetectable burst errors. This shows that the probability of undetectable burst errors can be modeled using a performability framework.

3. Modeling with Hash Functions

The CRC coding scheme essentially implements a hash function which converts error patterns into hash values, *e.g.*, the remainder from a polynomial division. For a message of length \( d \), if the length of a CRC code is \( m \), then there are totally \( 2^{d+m} \) error patterns (including the one with no error), as shown in Figure 1. This is the cardinality of \( S \). Moreover, for every \( 2^m \) consecutive error patterns, there is one error pattern that is “divisible”, *e.g.*, the remainder is 0. All error patterns that map onto the hash value of ‘0’ are the undetectable ones. Figure 1 depicts the hash function, detectable and undetectable error patterns, all within a performability framework [5].

![Figure 1: Hash Function and the Performability Framework](image-url)
The ratio of the number of undetectable error patterns to the total number of error patterns (including the one with no error) is equivalent to the divisible pattern (with remainder 0) to the cardinality of the set of all the remainders, i.e., $2^m$. Therefore, if the occurrence of every error pattern is equally likely, then the probability of undetectable burst errors is $2^{-m}$. Note that this attribute is purely dependent on the generator polynomial, not on the domain $S$.

When two CRC codes are applied in tandem, two hash functions $H_{CRC1}$ and $H_{CRC2}$ are used, as shown in Figure 2. Domain $S_1$ represents the set of the error patterns occurred in the highest $(d+m)$ bits, e.g., the data and the first code area; while domain $S_2$ represents the set of error patterns occurred in the entire message. The co-domain $R$ still consists of the two subsets, e.g., the detectable set and the undetectable set.

![Figure 2: Two Hash Functions and the Tandem CRCs Scheme](image)

4. Analysis

The decoding scheme can have four results: (1) both $H_{CRC1}$ and $H_{CRC2}$ detect the error (2) $H_{CRC1}$ detects the error but $H_{CRC2}$ does not detect the error (3) $H_{CRC1}$ does not detect the error but $H_{CRC2}$ detects the error (4) neither $H_{CRC1}$ nor $H_{CRC2}$ detects the error. Out of the four cases, only in case (4) where an error does not get detected.

As noticed in [4], the detection capability depends on the position of the error. Only the error patterns that cross the data and the $1^{st}$ CRC code areas, and the error patterns that cross the entire message can be undetectable. All other cases are detectable. That is, the highest bit of the error must exist in the data area, and the lowest bit must happen either in the $1^{st}$ CRC code area (this is the type-1 errors) or in the $2^{nd}$ CRC code area (this is referred to as the type-2 errors).

Denote an error pattern in the set $S_2$ as a polynomial $E(x)$, and an error pattern in the set $S_1$ as a polynomial $E_1(x)$. For a type-1 error, since the lowest $n$ bits in a message are error-free, $E(x)=E_1(x)X^n$. For a type-1 error to be undetectable, $E_1(x)$ must be divisible by $g(x)$ and $E(x)$ must be divisible by $h(x)$. Hence $E(x)=E_1(x)X^n = g(x)Q_1(x)X^n = h(x)Q_2(x)$. In other words, $E(x)$ must be divisible by $L(x)$, where $L(x)$ is the least common multiple polynomial of $g(x)$ and $h(x)$. The degree of $L(x)$ is $m+n-q$, where $q$ is the degree of the common factor of $g(x)$ and $h(x)$. If the two generator polynomials $g(x)$ and $h(x)$ are primal to each other, then $q$ is 0, and the probability of undetectable burst errors would be $1/2^{m+n}$. Note that $1/2^{m+n}=(1/2^m)\times (1/2^n)$, which implies that the event of an error to be undetectable using tandem CRCs scheme can be treated as two independent events. Note that the probability of undetectable burst errors is the product of the two probabilities of undetectable burst errors when applying a CRC scheme individually.
In the case of a type-2 error, the error can be divided into two parts, e.g., $E(x) = E_1(x)X^n + E_2(x)$. $E(x)$ is still divisible by $h(x)$ and $E_2(x)$ must be divisible by $g(x)$ when the error is not detectable. In [4], it was derived that $E_1(x)$ must not be divisible by $h(x)$ if the error is not detectable. To see this, note that the degree of $E_2(x)$ is less than $n$, hence $E_2(x)$ cannot be divisible by $h(x)$, because the degree of $h(x)$ is $n$. That is, although $E(x)$ is divisible by $h(x)$, $E_2(x)$ is not divisible by $h(x)$. Since $E(x) = E_1(x)X^n + E_2(x)$, it follows that $E_i(x)$ must also not be divisible by $h(x)$. Hence, $E_i(x)$ has the following relationship with $g(x)$ and $h(x)$:

$$E_i(x) = g(x)Q_i(x) = h(x)Q_{i0}(x) + R_{i0}(x), \text{ where } R_{i0}(x) \neq 0.$$

Due to the above relation, it seems like the two CRC schemes (hash functions) are related to each other. In particular, domain $S_i$ is dependent on the result of the second hash function ($H_{CRC2}$). In other words, if $E(x)$ passed the check using $h(x)$, then only those $E_i(x)$s that are not divisible by $h(x)$ are considered in the second check. However, recall that, from the previous analysis, that the probability of undetectable burst errors is purely dependent on the degree of the generator polynomial, but not on the domain of the hash function. Hence, the probability of undetectable burst errors of the tandem CRCs scheme can still be treated as the product of two independent events. As a result, the probability of undetectable burst errors is still in the order of $2^{(m+n)}$ assuming the two generator polynomials are primal to each other. This conclusion is consistent with the brief description presented in [2].

5. Conclusions

The probability of undetectable burst errors for systems using tandem CRCs scheme is analyzed in this paper under a performability framework. Moreover, it is realized that the domain of the hash function (which a CRC scheme implements) is irrelevant to the probability of the undetectable burst errors, and the generator polynomial determines the quantity of the attribute. This study concludes that tandem CRCs provides the capability to meet the stringent requirement on probability of undetectable burst errors.

References


Meng-Lai Yin is a full professor at the Department of Electrical and Computer Engineering, California State Polytechnic University, Pomona. She also works for Raytheon (previously Hughes Aircraft) Network Centric Systems since 1990. Her research interests include system analyses, parallel processing and quantum computing. She received her Ph.D. and M.S. degrees in Computer Science from University of California at Irvine, in 1995 and 1989, respectively.