Probability of Undetected Error with Redundant Data Transmission on a Binary Symmetric Channel without Memory

BÖRCSÖK J., HÖLZEL J., WACKER H. D.
Development
HIMA Paul Hildebrandt GmbH + Co KG
Albert-Bassermann-Strasse 28, D-68782 Bruehl
GERMANY
http://www.hima.com

Abstract: This paper investigates redundant data transmission on binary symmetric channels without memory protected by a linear code and the probability of undetected error. A simple formula suitable for numerical calculations is proved, improving a commonly used formula. A second formula for data transmission without cross check is given. The formula is applied to some frequently used CRC-16 polynomials with well known minimum distance to calculate block lengths maximal in order to achieve a specific Safety Integrity Level (SIL).

Key-Words: Linear Code, CRC, BSC, Probability Of Undetected Error, Bit Error Rate, Weight Distribution, Block length, Redundant Data Transmission, Cross Check, SIL

1 Introduction
Let there be given a binary symmetric channel without memory, a bit-error rate \( \varepsilon \) and a transmission procedure protected by a linear code C (i.e. a checksum procedure). Imagine for example a cyclic redundancy check CRC. Not the only one but a good measure for the performance of the code is the probability of undetected error (see [9]):

\[
(1) \quad p_{ue}(\varepsilon, C) = \sum_{l=1}^{n} A_l \varepsilon^l (1-\varepsilon)^{n-l}
\]

where

\( A_l = \text{weight distribution of } C \)
\( = \text{number of code words of weight } l \)
\( (\text{weight of a code word} = \text{number of bits equal to 1}) \)
\( \varepsilon = \text{bit error probability} \)
\( n = \text{block length} \)

In case of a poor (large) bit error probability a frequently used method to improve the performance of C is redundant (\( \mu \)-fold) data transmission together with cross check in the receiving device.

2 Data Protection by the Use of Linear Codes and \( \mu \)-fold Data Transmission
Each code word or block consists of a message to which a checksum is attached:

\[
c = (m_1, \ldots, m_k, s_0, \ldots, s_{r-1})
\]

Consider now a communication procedure transmitting each block twice and a linear code C performed separately on each of the two blocks. Further on, the receiving device is performing a cross check between both blocks (incl. both checksums). A block is accepted if only if there is no checksum fault and the two blocks inclusive checksum are identical.

Mathematically spoken this means that we defined a new Code \( C^{(2)} \) consisting of the code words

\[
c = (m_1, \ldots, m_k, s_0, s_1, m_1, \ldots, m_k, s_0, \ldots, s_{r-1})
\]

More generally: A \( \mu \)-fold transmission procedure together with \( \mu \)-fold protection by a checksum is characterized by a code \( C^{(\mu)} \) defined by its code words

\[
c = (m_1, \ldots, m_k, s_0, \ldots, s_{r-1}, \ldots, m_1, \ldots, m_k, s_0, \ldots, s_{r-1})
\]

that is

\[
(2) \quad C^{(\mu)} = \{ (x_1, \ldots, x_{\mu}) : x \in C \}
\]

\( C^{(\mu)} \) has to be carefully distinguished from the Cartesian product

\[
C^{(\mu)} = \{ (x_1, \ldots, x_{\mu}) : x_1, \ldots, x_{\mu} \in C \}
\]

The elements of \( C^{(\mu)} \) and \( C^{\mu} \) are typed in bold letters. The problem is now to find a relationship between

\[
p_{ue}(\varepsilon, C) \quad \text{and} \quad p_{ue}(\varepsilon, C^{(\mu)})
\]

A commonly used formula for the probability of undetected error with redundant transmission is given by (see [2]):

\[
(3) \quad p_{ue}(\varepsilon, C^{(\mu)}) \leq p_{ue}(\varepsilon, C)^{\mu}
\]

Normally deduced by heuristic arguments, equation (3) proves to be true. In section 3 we shall prove an exact formula for

\[
p_{ue}(\varepsilon, C^{(\mu)})
\]

improving and implicating (3). The new formula too will be suitable for numeric calculations.
3 The Probability of Undetected Error

3.1 The Main Result
At first let us state our main result: Theorem 1 will give a formula for $p_{ue}(\varepsilon, C^{(\mu)})$.

Theorem 1: The probability of undetected error of $C^{(\mu)}$ is given by

\begin{equation}
 p_{ue}(\varepsilon, C^{(\mu)}) = \sum_{l=1}^{n} A_l \varepsilon^{\mu l} (1-\varepsilon)^{\mu (n-l)}
\end{equation}

Proof: Let $x, y \in C$ be code words then, in the course of the proof, we shall use some notations:

$\mu y = \text{probability that } y \text{ is received, provided } x \text{ is sent}$

$d(x, y) = \text{Hamming distance between } x \text{ and } y = \text{number of bits in which } x \text{ and } y \text{ differ}$

Then, by the formula of the total probability, we get

\begin{equation}
 p_{ue}(\varepsilon, C^{(\mu)}) = \sum_{x \in C^{(\mu)}} (\sum_{y \in C^{(\mu)}} p(y|x)) p(x)
 = \sum_{x \in C} (\sum_{y \in C} p(y|x)^\mu) p(x)
 = \sum_{x \in C} \left( \sum_{y \in C} \varepsilon^{d(y,x)} (1-\varepsilon)^{\mu (n-d(y,x))} \right) \frac{1}{|C|}
 = \sum_{x \in C} \left( \sum_{y \in C} \varepsilon^{\mu d(y,x)} (1-\varepsilon)^{\mu (n-d(y,x))} \right) \frac{1}{|C|}
 = \sum_{x \in C} \left( \sum_{l=1}^{n} \varepsilon^{\mu l} (1-\varepsilon)^{\mu (n-l)} \right) \frac{1}{|C|}
 = \sum_{x \in C} \left( \sum_{l=1}^{n} A_l \varepsilon^{\mu l} (1-\varepsilon)^{\mu (n-l)} \right) \frac{1}{|C|}
\end{equation}

Corollary 1 states that the performance of the code $C^{(\mu)}$ at a bit-error rate $\varepsilon$ is at least as good as the performance of the code $C$ at a bit-error rate of $\varepsilon^\mu$.

Corollary 2: We have

\begin{equation}
 p_{ue}(\varepsilon, C^{(\mu)}) \leq p_{ue}(\varepsilon^\mu, C).
\end{equation}

Proof: By induction for $\mu = 1, 2, 3, \ldots$ we get

\begin{equation}
 (1-\varepsilon)^\mu \leq 1 - \varepsilon^\mu
\end{equation}

and therefore we have

\begin{equation}
 p_{ue}(\varepsilon, C^{(\mu)}) = \sum_{l=1}^{n} A_l \varepsilon^{\mu l} (1-\varepsilon)^{\mu (n-l)}
 \leq \sum_{l=1}^{n} A_l (\varepsilon^\mu)^l (1-\varepsilon)^{\mu (n-l)}
 = p_{ue}(\varepsilon^\mu, C).
\end{equation}

Corollary 2 is the well known result mentioned in section 2.

Corollary 3: We have

\begin{equation}
 p_{ue}(\varepsilon, C^{(\mu)}) \leq p_{ue}(\varepsilon, C)^\mu.
\end{equation}

Proof: Elementary calculus.

In the situation of Theorem 1 a code word $x \in C^{(\mu)}$ is sent, and the checksum procedure together with the cross check guarantee that the received $y$ again lies in $C^{(\mu)}$. What happens if we only check whether $y$ lies in the Cartesian product $C^{(\mu)}$? This means that only the checksums are verified and no cross check is done. One might expect, that (6) is true even without cross check. Unfortunately Theorem 4 states that this is not true.

Theorem 4: The probability of undetected error of $C^{(\mu)}$ is given by

\begin{equation}
 p_{ue}(\varepsilon, C^{(\mu)}) = (1-\varepsilon)^n + p_{ue}(\varepsilon, C)\mu - (1-\varepsilon)^{\mu n}.
\end{equation}

Proof: Similar to the proof of Theorem 1 by means of the multinomial theorem we get:

\begin{equation}
 p_{ue}(\varepsilon, C^{(\mu)}) = \sum_{x \in C^{(\mu)}} (\sum_{y \in C^{(\mu)}} p(y|x)) p(x)
 = \sum_{x \in C} \left( \sum_{y \in C} p(y|x)^\mu \right) p(x)
 = \sum_{x \in C} \left( \sum_{l=1}^{n} A_l \varepsilon^{\mu l} (1-\varepsilon)^{\mu (n-l)} \right) \frac{1}{|C|}
 = \sum_{x \in C} \left( \sum_{l=1}^{n} A_l \varepsilon^{\mu l} (1-\varepsilon)^{\mu (n-l)} \right) \frac{1}{|C|}
 = \sum_{x \in C} \left( \sum_{l=1}^{n} A_l \varepsilon^{\mu l} (1-\varepsilon)^{\mu (n-l)} \right) \frac{1}{|C|}
 = \sum_{x \in C} \left( \sum_{l=1}^{n} A_l \varepsilon^{\mu l} (1-\varepsilon)^{\mu (n-l)} \right) \frac{1}{|C|}
\end{equation}

In fact, (6) is not true for redundant transmission without cross check:
3.2 Safety Integrity Levels

Let us now have a closer look at data integrity according to IEC 68508 and analyze the effect of redundant transmission on maximal block lengths feasible for a specific Safety Integrity Level (SIL). Our calculations are based on the results about three CRC-16 \( C_i \), \( C_j \) and \( C_k \) generated by polynomials \( g_i \), \( g_j \) and \( g_k \) analyzed in [4].

\[
C_i : g_1 = x^{16} + x^{13} + x^{12} + x^{11} + \\
x^{10} + x^9 + x^8 + x^7 + x^5 + x^2 + 1 \\
C_j : g_3 = x^{16} + x^{14} + x^{12} + x^{11} + \\
x^8 + x^5 + x^3 + x + 1 \\
\]

and

\[
C_k : g_5 = x^{16} + x^{13} + x^{11} + x^9 + \\
x^7 + x^6 + x^5 + x^3 + x + 1
\]

\( C_i \) is optimal for a minimum distance (Hamming distance ) of \( d = 6 \), \( C_j \) is optimal for \( d = 5 \) and \( C_k \) is suitable for long block lengths. They are exemplary for a lot of other CRCs for which similar results are known. We did not check the rest of the CRC-16 treated in [4], because they are not proper for all block lengths, which means that \( p_{\text{ue}}(\varepsilon, C) \) is not an increasing function of \( \varepsilon \in [0,1/2] \). This means that a specific SIL being achieved for one \( \varepsilon \) could be violated for another smaller one, and more detailed inspections would be necessary. According to IEC 68508 the quantity \( \Lambda \) of undetected errors per hour is given by

\[
\Lambda = 3600 \cdot p_{\text{ue}}(\varepsilon, C) \cdot \nu \cdot (m - 1) \cdot 100
\]

where

\( \nu = \text{number of safety related messages per second} \)

\( m = \text{number of communicating devices} \)

100 = 1% - rule

For an example we decided to choose a relatively small \( \nu \) because for bigger \( \nu \) not all of the higher Safety Integrity Levels would be feasible. So for \( \nu = 1 \) and \( m = 1 \), we get

\[
\Lambda = 3.6 \cdot 10^5 \cdot p_{\text{ue}}(\varepsilon, C)
\]

If no details are known about the quality of the transmission especially about the electromagnetic compatibility (EMC) and nothing can be said about the bit error rate \( \varepsilon \), the German TÜV requires to do all calculations concerning \( \Lambda \) with \( \varepsilon = 10^{-2} \). Therefore for our analysis we took account of this bad value of the bit error rate. With the help of (8) and Theorem 1 the content of tables 1, 2 and 3 can be derived from the results in [4]. For our calculations we used the so called worst case formula

\[
p_{\text{ue}}(\varepsilon, C) = \sum_{i=0}^{n} \binom{n}{i} \varepsilon^i (1-\varepsilon)^{n-i},
\]

where \( d \) is the minimum distance of the CRC, and the results on \( d \) published in [4].

The tables below list the block lengths maximal in order to meet a specific Safety Integrity level (SIL) with \( \varepsilon = 10^{-2} \) and \( \varepsilon = 10^{-3} \). Columns 2 is taken from [8]. It contains the bounds on \( \Lambda \) for a specific Safety Integrity Level (SIL). Columns 3 and 4 contain the bounds on \( n \) for the single transmission mode respectively the double transmission. If the weight distribution of a code is completely known, better values of maximal block lengths are to be expected. Since the authors of [4] did not publish the weight distributions of their CRCs, we had to restrict our calculations to only making use of the minimum distances at different block lengths published in [4]. But for a demonstration of the effect of redundancy this should be sufficient.

### Table 1: Maximal block lengths for \( g_1 \)

<table>
<thead>
<tr>
<th>SIL</th>
<th>( \Lambda )</th>
<th>( n_{\text{max}} ) for single transmission</th>
<th>( n_{\text{max}} ) for single transmission</th>
</tr>
</thead>
<tbody>
<tr>
<td>high demand</td>
<td>( 10^{-8} )</td>
<td>22</td>
<td>151</td>
</tr>
<tr>
<td></td>
<td>( 10^{-7} )</td>
<td>22</td>
<td>151</td>
</tr>
<tr>
<td></td>
<td>( 10^{-6} )</td>
<td>22</td>
<td>151</td>
</tr>
<tr>
<td></td>
<td>( 10^{-5} )</td>
<td>23</td>
<td>151</td>
</tr>
</tbody>
</table>

### Table 2: Maximal block lengths for \( g_3 \)

<table>
<thead>
<tr>
<th>SIL</th>
<th>( \Lambda )</th>
<th>( n_{\text{max}} ) for single transmission</th>
<th>( n_{\text{max}} ) for double transmission</th>
</tr>
</thead>
<tbody>
<tr>
<td>high demand</td>
<td>( 10^{-8} )</td>
<td>23</td>
<td>247</td>
</tr>
<tr>
<td></td>
<td>( 10^{-7} )</td>
<td>24</td>
<td>257</td>
</tr>
<tr>
<td></td>
<td>( 10^{-6} )</td>
<td>26</td>
<td>257</td>
</tr>
<tr>
<td></td>
<td>( 10^{-5} )</td>
<td>26</td>
<td>257</td>
</tr>
</tbody>
</table>

### Table 3: Maximal block lengths for \( g_5 \)

<table>
<thead>
<tr>
<th>SIL</th>
<th>( \Lambda )</th>
<th>( n_{\text{max}} ) for single transmission</th>
<th>( n_{\text{max}} ) for double transmission</th>
</tr>
</thead>
<tbody>
<tr>
<td>high demand</td>
<td>( 10^{-8} )</td>
<td>22</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>( 10^{-7} )</td>
<td>27</td>
<td>126</td>
</tr>
<tr>
<td></td>
<td>( 10^{-6} )</td>
<td>27</td>
<td>211</td>
</tr>
<tr>
<td></td>
<td>( 10^{-5} )</td>
<td>27</td>
<td>353</td>
</tr>
</tbody>
</table>

More results with various CRCs about the size of the undetected error probability and their minimum distances as a function of the block length can be found
in [1], [3], [5], [6], [7], and [10]. With all these results, tables similar to those presented here, can be derived.

4 Conclusions
This paper contains two formulas for the probability of undetected error of redundant data transmission protected by a linear code on the binary symmetric channel without memory. A normally used formula is improved. Using results in [4], the effect of redundant transmission on maximal block lengths for achieving a specific Safety Integrity Level is investigated. The results are suitable for numerical calculations. They can be applied to CRCs with known minimum distances at different block lengths.

References: