Date: August 18, 1972

m (Div. & location ASDD Los Gatos Lab
.S. mail address):

Dept. & Bldg: A44/029
ielino & Tel. Ext.:

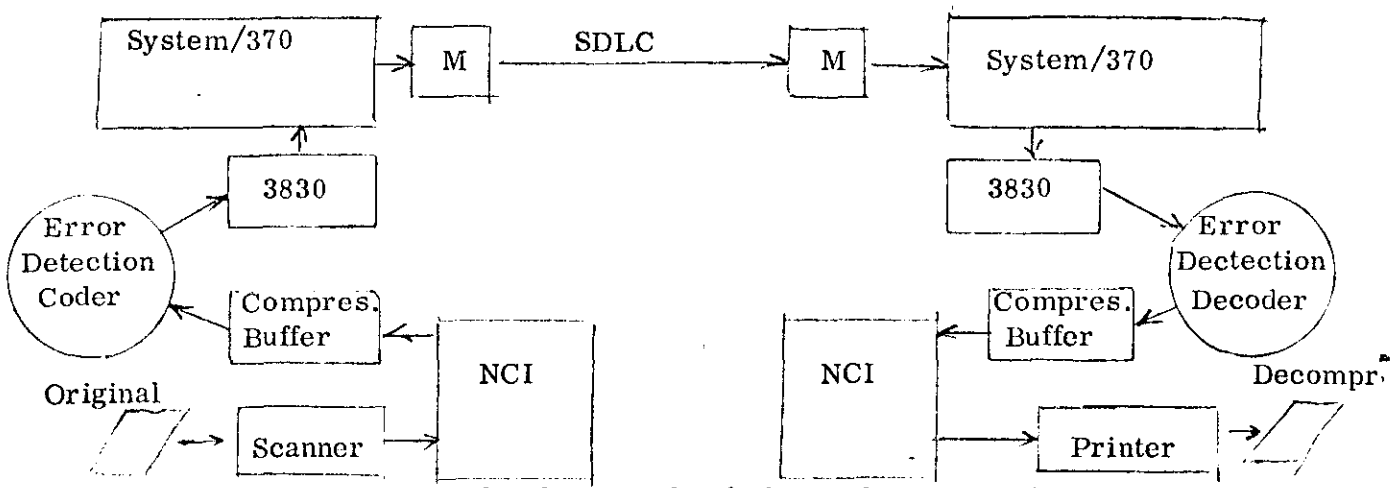
IBM CONFIDENTIAL

IBM

Subject: Error Detection Parameters for NCI Compressed Data Sets

Reference:

To: R. B. Arps



A preliminary examination has been made of what is feasible in the way error detection for compressed data in NCI.

Assumptions

Given an 8.5 X 11 inch document at 125 pels/inch has 1056 X 1364 = 1.44×10^6 bits per page. Depending upon the nature of the document, the compressed image will range from 150,000 bits to 500,000 bits/page.

What additional error control is needed to insure that the random undetected error bit rate is less than 10^{-13}? It is assumed that the maximum error protection features available in the 3830, System/370 and SDLC Communication links are used. The estimated level of error detection for the total system handling the compressed data is a probability of an undetected error bit of 10^{-9}.

The probability of a page going through the System with an undetected error ranges from one in 6600 pages for 150,000 bits per compressed page to one in 2,000 pages for 500,000 bits per compressed page.

The problem is to specify an error detecting code that can be inserted just after the compressor that will provide additional protection of 10^{-4} to improve the probability of undetected bit error from 10^{-9} to 10^{-13}.

In terms of number of pages per undetected error, this improvement would change the probability of a page being printed with an undetected error to the range of one in 20,000,000 to 66,000,000 pages.

Plan

I propose to recommend a cyclic error detection code as defined in Chapter 8 of Peterson, "Error-Correcting Codes (Second Edition)". Particular use is made of theorem 8.5 and 8.6 in Section 8.8 on error detection with cyclic code.

The detection of an error in the compressed data could be used to mark the page when printed with block line in the margin, by inverting the image after the point the error occured, or some other technique that would alert the human reader to the error.

The code should:

(1) detect random errors with less than $[\frac{10^{-13}}{10^{-9}}] = 10^{-4}$ probability if an undetected error, and

(2) must detect all the patterns that can slip through the various sub-systems.

a. The 3830 has an error correction feature based on the polynomial:

$$(X^3 + 1)(X^4 + 1) = X^7 + X^4 + X^3 + 1$$

An error patter, 10011001 could pass undetected in this section.

b. The System/370 has various parity error-detection, and error-correcting circuits. In the CPU of the Mod 145, a s/e correcting, d/e detecting code is used. Eight check bits are used with each double word of 64 bits. This is equivalent to a polynomial code of degree seven with an overall parity bit added.

c. In the SDLC data link the error detection is done by use of the polynomial:

$$(X+1)(X^{15} + X^{14} + X^{13} + X^{12} + X^4 + X^3 + X^2 + X + 1) =$$
$$X^{16} + X^{12} + X^5 + 1$$

This code is derived from root number 5427 of degree 15 on page 483 of Peterson (polynomial 170037E).

An error pattern of:10001000000100001 could slip through this section of the system.

Trial Design of Error Detection Code.

(1) Design for Random Errors.

We need a polynomial of degree $N \geq P$ where $P=\log_2 [\frac{10^{-9}}{10^{-13}}] = \log_2 10^4 = 3.32 \times 4 = 13.28$.

Therefore, N=14, which means at least 14 bit shift register is needed to obtain the protection against random errors specified. This criterion is derived from Peterson, Theorem 8.6. Since we know the existing error detection and correction features in the system change a random noise input

to a non-random distribution of errors, we must proceed to apply tests for detection of the individual errors that are more probable.

(2) Design for detection of Residual errors

a. 3830 Errors

To detect the pattern 10011001, we need at least an 8-bit feedback shift register. This is already exceeded by (1).

b. System/370 Errors.

The highest level of error correction is similar to the 3830 system. The conditions are already met by item (1).

c. SDLC Communication Link.

To detect the pattern 10001000000100001, we need at least a 17-bit feedback shift register.

(3) Design for Processing a Maximum Block of 500,000 Bits.

For N an integer, such that $N \geq \log_2 500,000$, we obtain $N1=19$, which means a minimum of a 19-bit shift register.

We next look in Marsh's Tables in Peterson to find a primitive ineducible polynomial of degree 19 and order of at least 500,000.

The value for j = 1 is 200047(OCTAL)F. The 'F' means this polynomial is primitive. The binary representation of this polynomial is:
10 000 000 000 100 111 giving the polynomial $X^{19} + X^5 + X^2 + X + 1$ as the candidate for our error detecting code.

Next, we compute the order (e) of this polynomial

$$e = \frac{2^{19} - 1}{GCD(1X524287, j)} = \frac{524287}{GCD\ (1X524287, 1)} = 524287.$$

Since $e > 500,000$, this polynomial should be satisfactory.

If the maximum compressed image is expected to be greater than (524287-19= 524268, we must look for the next higher value of m such that (2^m-1) is a prime number greater than the maximum image size. The next value after 19 is 31.

$2^{31} - 1 = 2,147,483,647$, which means a 31-bit feedback shift register. The j=1 root for m=31 is:

20, 000, 000011 (OCTAL)

Performance of the 19-bit Cyclic Code.

The cyclic code $X^{19} + x^5 + X^2 + 1$ can detect any burst error of 19 bits length or less. (th 8.5).

For error bursts of 20 bits length, the code can detect all but 2^{-18} of the possible patterns. The probability of an undetected random error burst of 20 bits length is, $p = 0.0000038$. For error bursts of 21 bits length or more, the code can detect all but 2^{-19} of the possible pattern. The probability of an undetected random error burst of 21 or greater length is $p = 0.000\ 0019$.

Consideration of Variable Block Length.

A five-bit counter plus detection circuits for count 16 could be used to find the end of page code 'FFFF' to stop the error detection logic in the variable block length mode of operation.

F. B. Wood

F. B. Wood

cc: S. Zell

attachment: reference list

References

F. B. Wood, IBM Los Gatos Report 16.186, December, 1970, "Evaluation of Error Detection Codes for 2400-band Digital Transmission over VHF Radio Line".

W. Wesley Peterson and E. J. Weldon, Jr., "Error Correction Codes, Second Edition", MIT Press (1972)

IBM Tele Processing Standards Sub-Committee on Coding. Archives Notes TPS-C-TN-3C, March 30, 1961.

Frederich F. Sellers, Jr., Mu-Yue Hsiao, Leroy W. Bearnson, "Error Detection Logic for Digital Computers", McGraw-Hill (1968)

Attachment to letter, F. B. Wood to R. B. Arps, dated August 18, 1972