Journal of ELECTRICAL ENGINEERING, VOL. 53, NO. 5-6, 2002, 138–142

COMMENTS ON "WEIGHT DISTRIBUTION OF SOME WEIGHTED SUM CODES FOR ERASURE CORRECTION"

Martin Rakús *

In [1], a new family of error detection codes called Weighted Sum Codes was proposed. In [2] it was noted that these codes are equivalent to lengthened Reed Solomon Codes or to a shortened version of lengthened Reed Solomon Codes over $GF(2^{(h/2)})$. It was also shown that it is possible to use these codes for error correction of one error in the codeword over $GF(2^{(h/2)})$. In [3], a weight spectra of Generalized Weighted Sum Codes (GWSC-II) constructed over $GF(8)$ were presented. In the present paper, some possibilities to improve erasure correction capabilities of these codes were investigated. At the end a brief analysis of erasure detection failure is given.

K e y w o r d s: weight spectra, erasure correction, GWSC, decoding failure

1 INTRODUCTION

Recently a family of generalized weighted sum codes (GWSC) for erasure correction was included into a new hybrid ARQ scheme for partially reliable transport services [4],[5]. In [6] it was shown that provision of reliable transport services in networks which experience a loss of packets can be very costly. Therefore correction of losses up to a certain level might be considered as an alternative approach. One of the major objectives when selecting an error control scheme for a real application is to keep redundancy as low as possible. GWSC offer efficient, low complexity software and hardware implementation as it is required in many high speed networks. In [7], the weight spectra of GWSC codes for correction of up to 3 erasures were analyzed. It was shown that the GWSC can reliably correct up to 3 erasures in a codeword. Exact weight spectra of GWSC-II were presented in [3]. In this paper, two possibilities of improving erasure correction capabilities of GWSC-II will be described.

2 WEIGHTED SUM CODES

A new family of codes that use polynomial arithmetics and process r bit symbols was introduced in [1].

The next part is a shortened description of these codes. The symbols $Q_1, Q_2, \cdots, Q_{q-2}, P_0, P_1, \cdots, P_{z-1}$ are binary strings of length r and each Q_i and P_i can be regarded as a member of $GF(q)$ where $q = 2^r$. α is a primitive element of the finite field $GF(q)$ and z is the number of erasures in a codeword.

$Q_{q-2}, Q_{q-3}, \cdots, Q_1$ is a message consisting of $q-2$ symbols from $GF(2^r)$ and it is encoded as a string of n symbols, where n is a codeword length, by adding the check symbols $P_0, P_1, \cdots, P_{z-1}$. Check symbols are calculated by the following rules:

a) In the first class of Generalized Weighted Sum Codes denoted as GWSC-I, the parity symbols are evaluated as follows:

$$
\begin{aligned}
P_0 &= P_{z-1} \oplus P_{z-2} \oplus \cdots \oplus P_1 \oplus \sum_{i=1}^{q-2} Q_i, \\
P_1 &= P_{z-1} \oplus \cdots \oplus P_2 \oplus \sum_{i=1}^{q-2} \alpha^i \otimes Q_i, \\
P_2 &= P_{z-1} \oplus \cdots \oplus P_3 \oplus \sum_{i=1}^{q-2} \alpha^{2i} \otimes Q_i, \\
&\vdots \\
P_{z-1} &= \sum_{i=1}^{q-2} \alpha^{(z-1)i} \otimes Q_i. \qquad (1)
\end{aligned}
$$

It can be shown that the first type of Generalized Weighted Sum Codes is equivalent to codes with the **H** matrix given by (3).

b) In the second class of Generalized Weighted Sum Codes denoted as GWSC-II the parity symbols are evaluated as follows:

$$
\begin{aligned}
P_0 &= \sum_{i=1}^{q-2} Q_i, \\
P_1 &= \sum_{i=1}^{q-2} \alpha^i \otimes Q_i, \\
P_2 &= \sum_{i=1}^{q-2} \alpha^{2i} \otimes Q_i, \\
&\vdots \\
P_{z-1} &= \sum_{i=1}^{q-2} \alpha^{(z-1)i} \otimes Q_i. \qquad (2)
\end{aligned}
$$

* Slovak University of Technology, Faculty of Electrical Engineering and Information Technology, Department of Telecommunications, Ilkovičova 3, 812 19 Bratislava, Slovakia, E-mail: rakus@ktl.elf.stuba.sk

ISSN 1335-3632

$$H=\begin{pmatrix} 1 & 1 & \cdots & 1 & 1 & 1 & \cdots & \cdots & 1\\ \alpha^{(q-2)} & \alpha^{(q-3)} & \cdots & \alpha & 1 & \cdots & \cdots & 1 & 0\\ \alpha^{2(q-2)} & \alpha^{2(q-3)} & \cdots & \alpha^2 & 1 & \cdots & 1 & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ \alpha^{(z-1)(q-2)} & \alpha^{(z-1)(q-3)} & \cdots & \alpha^{(z-1)} & 1 & 0 & \cdots & \cdots & 0 \end{pmatrix} \tag{3}$$

$$H=\begin{pmatrix} 1 & 1 & \cdots & 1 & 1 & 0 & 0 & 0 & \cdots & \cdots & 0\\ \alpha^{(q-2)} & \alpha^{(q-3)} & \cdots & \alpha & 0 & 1 & 0 & 0 & \cdots & \cdots & 0\\ \alpha^{2(q-2)} & \alpha^{2(q-3)} & \cdots & \alpha^2 & 0 & 0 & 1 & 0 & \cdots & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ \alpha^{(z-1)(q-2)} & \alpha^{(z-1)(q-3)} & \cdots & \alpha^{(z-1)} & 0 & 0 & 0 & 0 & \cdots & \cdots & 1 \end{pmatrix} \tag{4}$$

It can be shown that the second type of Generalized Weighted Sum Codes is equivalent to the codes with **H** matrix given by (4).

Erasure correction decoding for both groups of GWSC codes are described in [8].

For $z \geq 4$ these codes do not guarantee the correction of z erasure for all codewords, which can be seen from the following weight spectra of GWSC-II codes

Number of parity symbols $z = 1$:

$$a_0 = 1,\ a_2 = 147,\ a_3 = 1470,\ a_4 = 10535,$$
$$a_5 = 44100,\ \ a_6 = 102949,\ a_7 = 102942$$

Number of parity symbols $z = 2$:

$$a_0 = 1,\ a_3 = 392,\ a_4 = 2450,\ a_5 = 14896,$$
$$a_6 = 51352,\ a_7 = 102984,\ a_8 = 90069$$

Number of parity symbols $z = 3$:

$$a_0 = 1,\ a_4 = 882,\ a_5 = 3528,\ a_6 = 19992,$$
$$a_7 = 57456,\ a_8 = 101493,\ a_9 = 78792$$

Number of parity symbols $z = 4$:

$$a_0 = 1,\ a_4 = 56,\ a_5 = 1428,$$
$$a_6 = 5250,\ a_7 = 24920,\ a_8 = 62895,$$
$$a_9 = 98644,\ a_{10} = 68950$$

Number of parity symbols $z = 5$:

$$a_0 = 1,\ a_5 = 154,\ a_6 = 2310,$$
$$a_7 = 6930,\ a_8 = 30415,\ a_9 = 66960,$$
$$a_{10} = 95018,\ a_{11} = 60326$$

Number of parity symbols $z = 6$:

$$a_0 = 1,\ a_6 = 350,\ a_7 = 3444,$$
$$a_8 = 8715,\ a_9 = 36120,\ a_{10} = 69930,$$
$$a_{11} = 90804,\ a_{12} = 52780$$

Number of parity symbols $z = 7$:

$$a_0 = 1,\ a_7 = 686,\ a_8 = 4893,$$
$$a_9 = 10290,\ a_{10} = 42336,\ a_{11} = 71442,$$
$$a_{12} = 86338,\ a_{13} = 46158$$

Number of parity symbols $z = 8$:

$$a_0 = 1,\ a_7 = 336,\ a_8 = 2135,$$
$$a_9 = 6468,\ a_{10} = 16506,\ a_{11} = 43848,$$
$$a_{12} = 66976,\ a_{13} = 79716,\ a_{14} = 46158.$$

Weight spectra of GWSC-I codes exhibit the same property as weight spectra of GWSC-II codes with regard to erasure correction capabilities. From the presented spectra it can be seen that if the number of parity symbols is equal to z then the code will fail to correct z erasures in all cases. For example in the code given by (4) and for $z = 4$ there are 56 codewords with weight four. The number of such codewords is relatively small in comparison with the number of all codewords (262144) for that particular code. Nevertheless, if we allow to use more parity symbols than z, we can for some cases guarantee correction of all combinations with z erasures. For example the code with $z = 5$ does not contain codewords with weight 4 and thus can guarantee correction of 4 erasures.

3 POSSIBLE IMPROVEMENTS OF ERASURE CORRECTION CAPABILITIES OF GWSC–II CODES

From the previous section it can be concluded that GWSC codes can not guarantee correction of z erasures in a codeword for all possible instances. Therefore some attempts were made to explore a possible improvements. Two methods were investigated:

- combination of parity check columns vectors of **G**
- adding of a new column to **H**

For comparison of both methods, weight spectra are presented.

3.1 Combination of parity check columns vectors of G

As it could be seen from the presented weight spectra of GWSC-II codes, for $z \geq 4$ these codes are not able to correct all z erasure for all cases. So as the first attempt to eliminate this drawback, a weight spectra for all possible combinations of parity-check columns of **G** were calculated. Parity submatrix of generating matrix **G** of GWSC-II codes consists of some columns of Vandermond matrix. So all columns of such a matrix can be created as powers of the second column which have in $GF(8)$ form:

$$\boldsymbol{c}_2^\top = \left(\alpha^6, \alpha^5, \alpha^4, \alpha^3, \alpha^2, \alpha\right) \tag{5}$$

All other parity columns can be generated as a powers of (5): $\boldsymbol{c}_i = \boldsymbol{c}_2^i$, $i = 0, \cdots, q-2$, where $\boldsymbol{c}_i$ is the i-th parity column of **G**. The number of all possible combinations for parity columns of **G**, denoted as $comb_z(q-1)$, is

$$comb_z(q-1) = \binom{q-1}{z} \tag{6}$$

where q is the number of field elements of $GF(q)$ and z is the number of erasures. Calculated weight spectra (see Tab. 1) for all combinations for $z = 4, 5, 6$ were presented in [9]. Single codes are numbered the way which corresponds to the powers of parity columns. For example for $z = 4$ a code denoted as 0123 has a **G** matrix of the form:

$$\boldsymbol{G} = (\boldsymbol{I} \mid \boldsymbol{c}_2^0\, \boldsymbol{c}_2^1\, \boldsymbol{c}_2^2,\, \boldsymbol{c}_2^3). \tag{7}$$

By comparing the selected results from [9] (see Tab.1) it can be seen that in the case of $z = 4$ certain combinations exist for which $a_4 = 49$ (*eg* code with the combination number 0124) which is a lower value than that of the original code. This results in a lower probability of decoding failure, as will be analyzed later. The second conclusion is that for $z = 5, 6$ the combinations of powers of parity columns do not have the effect on the code weight spectrum.

Table 1. Comparison of the results selected from [9]

code	0124	0123	2346	0346	1346
a_0	1	1	1	1	1
a_1	0	0	0	0	0
a_2	0	0	0	0	0
a_3	0	0	0	0	0
a_4	49	56	70	84	98
a_5	1 470	1 428	1 344	1 260	1 232
a_6	5 145	5 250	5 460	5 670	5 600
a_7	25 060	24 920	24 640	24 360	24 640
a_8	62 790	62 895	63 105	63 315	62 965
a_9	98 686	98 644	98 560	98 476	98 672
a_{10}	68 943	68 950	68 964	68 978	68 936

3.2 Adding a new column to H

The second method, in the attempt to improove erasure correction capabilities of GWSC-II codes, was adding a new column to control matrix **H**. This idea was suggested by Ernst Gabidulin from the Moscow Institute of Physics and Technology. In this case the number of information symbols k is risen by one to: $k+1$, thus $k = q-1$. Again, information symbols $Q_0, Q_1, \cdots, Q_{q-2}$ are binary strings of lenght r. A message consisting of $k+1$ information symbols is encoded by adding parity symbols $P_0, P_1, \cdots, P_{z-1}$. Parity symbols for GWSC-II codes are evaluated as follows.

$$\begin{aligned} P_0 &= \sum_{i=0}^{q-2} Q_i, \\ P_1 &= \sum_{i=0}^{q-2} \alpha^i \otimes Q_i, \\ P_2 &= \sum_{i=0}^{q-2} \alpha^{2i} \otimes Q_i, \\ &\vdots \\ P_{z-1} &= \sum_{i=1}^{q-2} \alpha^{(z-1)i} \otimes Q_i. \end{aligned} \tag{8}$$

The weight spectra of such a modified GWSC-II codes follow.
Number of parity symbols $z = 1$:

$$a_0 = 1,\ a_2 = 196,\ a_3 = 2352,$$
$$a_4 = 21070,\ a_5 = 117600,\ a_6 = 411796,$$
$$a_7 = 823536,\ a_8 = 720601$$

Number of parity symbols $z = 2$:

$$a_0 = 1,\ a_3 = 588,\ a_4 = 4410,$$
$$a_5 = 33516,\ a_6 = 154056,\ a_7 = 463428,$$
$$a_8 = 810621,\ a_9 = 630532$$

Number of parity symbols $z = 3$:

$$a_0 = 1,\ a_4 = 1470,\ a_5 = 7056,$$
$$a_6 = 49980,\ a_7 = 191520,\ a_8 = 507465,$$
$$a_9 = 787920, a_{10} = 551740$$

Number of parity symbols $z = 4$:

$$a_0 = 1,\ a_4 = 98,\ a_5 = 2548,$$
$$a_6 = 11760,\ a_7 = 68180,\ a_8 = 230965,$$
$$a_9 = 542332,\ a_{10} = 758520,\ a_{11} = 482748$$

Number of parity symbols $z = 5$:

$$a_0 = 1,\ a_5 = 294,\ a_6 = 4410,$$
$$a_7 = 17262,\ a_8 = 90195,\ a_9 = 269010,$$
$$a_{10} = 569478,\ a_{11} = 724122,\ a_{12} = 422380$$

Number of parity symbols $z = 6$:

$$a_0 = 1,\ a_6 = 686,\ a_7 = 7210,$$
$$a_8 = 23415,\ a_9 = 116130,\ a_{10} = 304290,$$
$$a_{11} = 589470,\ a_{12} = 686392,\ a_{13} = 369558$$

Number of parity symbols $z = 7$:

$$a_0 = 1,\ a_7 = 1372,\ a_8 = 11417,$$
$$a_9 = 28812,\ a_{10} = 148176,\ a_{11} = 333396,$$
$$a_{12} = 604366,\ a_{13} = 646212,\ a_{14} = 323400$$

Number of parity symbols $z = 8$:

$$a_0 = 1,\ a_7 = 686,\ a_8 = 5579,\ a_9 = 16814,$$
$$a_{10} = 60858,\ a_{11} = 177282,\ a_{12} = 348292,$$
$$a_{13} = 564186,\ a_{14} = 600054,\ a_{15} = 323400$$

From the presented weight spectra it can be seen that the modified code is able reliably to detect and correct up to 3 erasure in codeword. For $z \geq 4$ this code is not able to correct reliably erasure in all cases. Nevertheless the code rate for modified GWSC-II code is

$$R = \frac{(q-1)}{(q-1+z)} \tag{9}$$

which is higher than for ordinary GWSC-II code. Comparison of code rates for ordinary and modified GWSC-II codes can be found in [9].

4 ANALYSIS OF ERASURE CORRECTION FAILURE OF GWSC CODES

Decoding of GWSC codes was described in [8]. The exact values of erasure can be calculated from the linear system of control equations. In the next, a brief description of erasure decoding for two erasure will follow. Positions a and b of symbol erasure Z_1, Z_2 in the received word

$$\hat{\hat{c}} = (\hat{Q}_{q-2}, \hat{Q}_{q-3}, \cdots, \hat{Q}_1, \hat{P}_1, \hat{P}_0) \tag{10}$$

a and b are known. For example the erasure Z_1 is on position a and the erasure Z_2 on the position b. The values of symbol erasure W_a and W_b on corresponding positions are known.

Decoding with correction of two symbol erasure in one codeword over $GF(q)$ can be made in the following steps

1) evaluation of syndromes S_1 and S_0:

$$S_0 = \hat{P}_0 \oplus \sum_{i=1}^{q-2} \hat{Q}_i, \tag{11}$$

$$S_1 = \hat{P}_1 \oplus \sum_{i=1}^{q-2} \alpha^i \otimes \hat{Q}_i \tag{12}$$

If the syndromes are zero: end of decoding, else go to 2.

2) evaluation of the erasure values from the control equations:

2a) if exactly two erasure on the position a and b occur, then:

$$S_0 \oplus Z_1 \oplus Z_2 = 0 \tag{13}$$

$$S_1 \oplus \alpha^a \otimes Z_1 \oplus \alpha^b \otimes Z_2 = 0 \tag{14}$$

hence

$$Z_2 = \frac{S_1 \oplus S_0 \otimes \alpha^a}{\alpha^a \oplus \alpha^b} \tag{15}$$

and

$$Z_1 = \frac{S_1 \oplus S_0 \otimes \alpha^a}{\alpha^a \oplus \alpha^b} \oplus S_0 \tag{16}$$

For $z = 2$ the minimum non–zero weight is a_3, therefore the system of linear equations (13),(14) has always a unique nontrivial solution. For codes where exists minimum non–zero weight codewords corresponding with the number of erasure *eg* for $z = 4$, the minimum non–zero weight is a_4. In such a case when erasure are exactly at the same positions as non–zero symbols in codewords with weight a_4, two events can occur:

a) erasure values denoted as: $W_a, W_b, W_c, W_d = 0$ – in such a case all syndromes: $S_0, \cdots, S_3 = 0$, which leads to incorrect decoding

b) if erasure values $W_a, W_b, W_c, W_d \neq 0$, the system of linear control equations becomes linearly dependent and does not have a unique nontrivial solution.

Therefore the probability of decoding failure for $z \geq 4$ depends on the number of codewords with minimal weight. If the number of codewords with minimal non–zero weight is denoted as a_z, n is the codeword length, z is the number of erasure and q is the size of an used field, the probability of decoding failure denoted as P_{ED_Z} can be calculated as

$$P_{ED_z} = \frac{a_z}{\binom{n}{z})(q-1)^z} \tag{17}$$

For ordinary GWSC-II codes constructed over $GF(8)$ calculation of P_{ED_z} will follow

$$\begin{array}{lll} z=4 & a_4 = \ \ 56 & P_{ED_{z=4}} \doteq 1.11065 \times 10^{-4} \\ z=5 & a_5 = 154 & P_{ED_{z=5}} \doteq 1.98330 \times 10^{-5} \\ z=6 & a_6 = 350 & P_{ED_{z=6}} \doteq 3.21964 \times 10^{-6} \\ z=7 & a_7 = 686 & P_{ED_{z=7}} \doteq 4.85423 \times 10^{-7} \end{array} \tag{18}$$

For codes described in subsection 3.1 the probability of decoding failure was calculated only for $z = 4$ because for $z > 4$ the combination of parity columns did not change the weight distribution. Calculation of $P_{ED_{z=4}}$ according Tab.1 follows:

$$\begin{aligned} a_4 = 49 \quad & P_{ED_{z=4}} \doteq 9.71817 \times 10^{-5} \\ a_4 = 56 \quad & P_{ED_{z=4}} \doteq 1.11065 \times 10^{-4} \\ a_4 = 70 \quad & P_{ED_{z=4}} \doteq 1.38831 \times 10^{-4} \\ a_4 = 84 \quad & P_{ED_{z=4}} \doteq 1.66597 \times 10^{-4} \\ a_4 = 98 \quad & P_{ED_{z=4}} \doteq 1.94363 \times 10^{-4} \end{aligned} \tag{19}$$

For codes described in subsection 3.2 the probability of decoding failure was calculated as follow:

$$\begin{aligned} z = 4 \quad a_4 = & 98 \quad P_{ED_{z=4}} \doteq 1.23686 \times 10^{-4} \\ z = 5 \quad a_5 = & 294 \quad P_{ED_{z=5}} \doteq 2.20868 \times 10^{-5} \\ z = 6 \quad a_6 = & 686 \quad P_{ED_{z=6}} \doteq 3.39796 \times 10^{-6} \\ z = 7 \quad a_7 = & 1372 \quad P_{ED_{z=7}} \doteq 4.85423 \times 10^{-7} \end{aligned} \tag{20}$$

By comparing the results from (19) it is possible to conclude that just by picking a better combination of parity columns (in this case the combination 0124) it is possible to reduce the probability of decoding failure from $P_{ED_{z=4}} \doteq 1.11065 \times 10^{-4}$ to $P_{ED_{z=4}} \doteq 9.71817 \times 10^{-5}$ which represents approximately 12.5%. By comparing the results from (18) and (20) it is possible to conclude that by adding one information symbol we can increase the code rate while the P_{ED_Z} is just slightly increased.

5 CONCLUSION

It was shown that by introducing simple modifications it is possible to improve some properties of Generalized Weighted Sum Codes. Exact weight spectra of both modifications were presented. From the analysis of the probability of decoding failure it is possible to conclude that these codes could be effectively used in applications where partially reliable services are allowable.

References

[1] McAULEY, A. J.: Weighted Sum Codes for Error Detection and their Comparison with Existing Codes, IEEE Trans. on Networking **2** No. 1 (1994), 16-22.

[2] FARKAŠ, P.: Comments on "Weighted Sum Codes for Error Detection and their Comparison with Existing Codes", IEEE Trans. on Networking **3** No. 2 (1995), 129-130.

[3] FARKAŠ, P.—RAKÚS, M.: Weight Distribution of Some Weighted Sum Codes for Erasure Correction, TSP'99, 22^{nd} Int. Conference on Telecommunications and Signal Processing, Brno, Czech Republic, September 1999, pp. 106-109.

[4] FARKAŠ, P.—ZEPERNICK, H. J.—WYSOCKI, T.: A Hybrid ARQ Scheme for Partially Reliable Transport Services, Proc. IWSSIP'99, Bratislava, June 1999..

[5] FARKAŠ, P.—ZEPERNICK, H. J.—WYSOCKI, T.: Performance of a Hybrid ARQ Scheme for Partially Reliable Transport Services, Proc. VTC 1999 - Fall, Amsterdam, September 1999.

[6] MARASALI, R. Amer,—P. B.—CONRAD, P. T.: Retransmission – Based Partially Reliable Transport Service: "An analytic model", Proc. 1996, pp. 5c.4.1-5c.4.9.

[7] FARKAŠ, P.—RAKÚS, M.: Weight Spectra of Two and Three Erasure Correcting Weighted Sum Codes over GF(8)., DSP'99, The 4^{th} Int. Conference on Digital Signal Processing, Herľany, Slovakia, September 1999, pp. 1-3.

[8] FARKAŠ, P.: Generalized Weighted Sum Codes for Erasure Correction, ERK'98, Electrotechnical and Computer Science Conference, Portorož, Slovenija, September 1998, pp. 21-24.

[9] RAKÚS, M.: Analysis of Correction Capabilities of the New Error Correcting Codes Using Computer, Diploma Thesis, Slovak University of Technology, Faculty of Electrical Engineering and Information Technology, Department of Telecommunications, Bratislava, Slovakia, Jan 2001.

Received 11 December 2001

Martin Rakús (Ing) graduated in radioelectronics from the Slovak University of Technology in 2001. Since 1995 he has been with the Department of Telecommunications. Currently he works at the same department as an Assistant Professor. Since 2001 he has been a PhD student. His primary research interests are error control coding and communication systems.