Journal of ELECTRICAL ENGINEERING, VOL 70 (2019), NO3, 256–258

sciendo

COMMUNICATIONS

Erasure decoding of five times extended Reed-Solomon codes

Martin Rakús[*], Peter Farkaš[*,], Tomáš Páleník[*]**

Recently a new family of error control codes was proposed which are equivalent to five times extended Reed-Solomon codes. In this paper an erasure decoding algorithm for these codes is proposed.

Keywords: extended Reed-Solomon codes, decoding, algorithm, erasures

1 Introduction

Reed Solomon (RS) codes [1] are at present the most frequently used error control codes in practice [2]. They are used for example in such standards as DVB-T or in CD applications [3]. From a theoretical point of view RS codes belong to linear block codes and could be described the same way as cyclic codes are [4]. A linear block code is defined as a k-dimensional subspace of an n-dimensional vector space constructed over a finite field $GF(q)$. It is often described using a triple $[n, k, d_m]$ in which n is the codeword length, k is the number of information symbols in each codeword and d_m is the code distance, which is a minimal Hamming distance between any two codewords from the code. The Hamming distance is defined as the number of coordinates (symbols) by which two codewords (or in general two vectors) differ. The code distance and the number of correctable errors (denoted as t) in a linear block code is linked by the following inequality

$$d_m \geq 2t + 1 . \tag{1}$$

A convenient way to specify a linear block code is to use matrix notation. One such matrix is the control matrix $\boldsymbol{H}$.

Recently in [5] a new family of error control codes constructed over $GF(q)$ was proposed, where $q = 2^m$ and m is an odd integer, using control matrix

$$\boldsymbol{H} = \begin{bmatrix} \alpha^0 & \alpha^0 & \dots & \alpha^0 & \alpha^0 & \alpha^0 & 1 & 0 & 0 & 0 & 0 \\ \alpha^{(q-2)} & \alpha^{(q-3)} & \dots & \alpha^2 & \alpha^1 & \alpha^0 & 0 & 1 & 0 & 0 & 0 \\ \alpha^{2(q-2)} & \alpha^{2(q-3)} & \dots & \alpha^4 & \alpha^2 & \alpha^0 & 0 & 0 & 1 & 0 & 0 \\ \alpha^{3(q-2)} & \alpha^{3(q-3)} & \dots & \alpha^6 & \alpha^3 & \alpha^0 & 0 & 0 & 0 & 1 & 0 \\ \alpha^{4(q-2)} & \alpha^{4(q-3)} & \dots & \alpha^8 & \alpha^4 & \alpha^0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} . \tag{2}$$

This infinite family of codes can be characterized by the following triple $[n = q + 4, q - 1, 5]$. In [5] the construction of these codes together with the proof that each code from this family has $d_m = 5$ was presented. In [5] no decoding method was described. However, to make these codes useful in practice, knowing an implementable decoding method is necessary. Therefore, in this short communication a new decoding algorithm for erasure corrections for these codes is proposed.

2 Some notes on RS code decoding

As was already mentioned, RS codes have a broad range of applications [2]. Consequently, they have long been in the focus of coding theorists as well as coding practitioners [3]. There is vigorous research concerning these codes and their decoding algorithms even 60 years after their discovery, which could be documented by the following selected references [6–12]. Therefore, there are numerous known algorithms for their encoding as well as for decoding.

In this paper we will concentrate only on a subset of such algorithms, namely the syndrome methods which are relevant to the proposed algorithm for the five times extended RS codes.

The main practical motivation for using error control codes is to decrease the influence of impairments which can occur during information transmission or storage. Usually the impairments which could be handled efficiently by RS codes are categorized as errors or erasures. The errors in an RS code codeword are symbol errors and each such symbol error could be described by two unknowns X and Y. For example, the i-th error is determined by its error value Y_i and by its position, which is given by the corresponding error locator X_i.

Note. We will restrict our attention to finite fields with characteristics two; therefore we will suppose that both X_i and Y_i are elements from $GF(2^m)$, where $m > 1$ is an odd integer.

* The Institute of Multimedia Information and Communication Technologies, Faculty of Electrical Engineering and Information Technology, Slovak University of Technology, Ilkovičova 3, 812 19 Bratislava, Slovakia, rakus@ut.fei.stuba.sk, palenik@ut.fei.stuba.sk, ** The Institute of Applied Informatics, Pan-European University, FI PEVŠ Tematínska 10, 851 05 Bratislava Slovakia,p.farkas@ieee.org

DOI: 10.2478/jee-2019–0035, Print (till 2015) ISSN 1335-3632, On-line ISSN 1339-309X

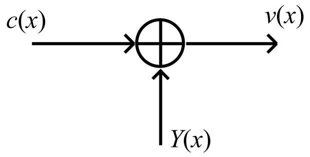

Fig. 1. Additive error channel model

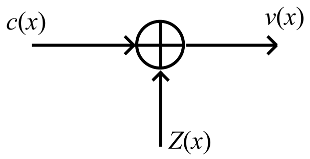

Fig. 2. Erasure channel model

To correct one error, the decoder needs to calculate both values for this error. The occurrence of errors in a codeword caused by transmission or writing/reading from a storage system could be modeled using an additive channel as is depicted in Fig. 1.

In Fig. 1 and Fig. 2 $\oplus$ denotes an addition of two vectors over $GF(2^m)$, $c(x)$ is the transmitted codeword, $Y(x)$ is an error polynomial, $Z(x)$ is an erasure polynomial and $v(x)$ is the received polynomial, which can contain errors or erasures.

On the other hand, erasure can be described by a single unknown, namely by its value denoted for the i-th erasure as Z_i. The position of the erasure is known to the decoder before the decoding starts. In practice this happens for example when the symbol in the codeword at the known position is missing. To correct one erasure the decoder must calculate only the value of the unknown erasure Z_i and then add it to the known position of the corresponding erasure as is shown in Fig. 2.

The most common algorithms for RS code decoding could be from a high-level point of view described as a solution of system of equations constructed over finite fields.

Before these equations could be formed it is necessary to calculate the syndrome values. In order to calculate these syndrome values $2t$ roots are inserted into the received polynomial.

$$v(x) = v_{n-1}x^{n-1} + v_{n-2}x^{n-2} + \cdots + v_1x^1 + v_0x^0. \quad (3)$$

For example, if the set of $2t$ consecutive roots starts with α^0 we will get the following syndrome values

$$\begin{aligned} S_0 &= v(\alpha^0), \\ S_1 &= v(\alpha^1), \\ S_2 &= v(\alpha^2), \\ &\vdots \\ S_{2t-1} &= v(\alpha^{2t-1}). \end{aligned} \quad (4)$$

For erasure decoding, a system of linear equations could be used, which contains erasures as unknown and syndromes as constants

$$S_k = \sum_{i=1}^{\zeta} Z_i X_i^k; \quad i = 1, 2, \ldots, \zeta, \quad (5)$$

where it is assumed that the number of actually occurring erasures denoted as ζ is smaller or equal to the number of correctable erasures denoted as z. The maximal number of correctable erasures z in each codeword is connected with the code distance by the following relationship

$$z + 1 \leq d_m. \quad (6)$$

It is obvious that the number of linearly independent equations in (5) needs to be at least ζ or expressed with other words – for one erasure correction one linearly independent equation is necessary in (5).

3 One algorithm for erasure correction of five times extended RS codes

In some situations, the possibility of correcting erasures in the received information can be advantageous. As was already mentioned, the erasures can have different causes. For example, the corresponding symbols can be lost during the transmission or the detector can delete the least reliable symbols and give the decoder the additional information of which symbols were deleted. Since the code distance of the analyzed codes $d_m = 5$, the new codes from [5] can correct up to 4 erasures in one codeword or, mathematically expressed, $z = 4$. In this section we will describe one method of correcting 4 symbol erasures. Their values will be denoted as Z_a, Z_b, Z_c and Z_d.

The new codes are equivalent to five times extended RS codes, therefore similar methods could be used for their decoding. In contrast to ordinary RS codes, five times extended RS codes contain five additional symbols. Therefore, to clearly highlight the differences in decoding we will use the following vector notation for a codeword

$$\boldsymbol{c} = (c_{q-2}, c_{q-3}, \ldots, c_0, p_4, p_3, p_2, p_1, p_0). \quad (7)$$

The receiver receives a vector

$$\mathbf{v} = (v_{q+3}, v_{q+2}, \ldots, v_4, v_3, v_2, v_1, v_0), \quad (8)$$

with potentially corrupted received versions of symbols c_i and p_i which we denote as $\hat{c}_i$ and $\hat{p}_i$, where c_i, p_i, $\hat{c}_i$ and $\hat{p}_i$ are elements of $GF(2^m)$. Using this notation, the received vector could also be expressed as follows

$$\mathbf{v} = (\hat{c}_{q-2}, \hat{c}_{q-3}, \ldots, \hat{c}_0, \hat{p}_4, \hat{p}_3, \hat{p}_2, \hat{p}_1, \hat{p}_0). \quad (9)$$

In (9), the erasure positions are known. Erasures: Z_a, Z_b, Z_c and Z_d are in the corresponding positions which are denoted as a, b, c, d. The values of received symbols at erasure (known) positions of codewords are denoted as: v_a, v_b, v_c and v_d. The original values c_a, c_b, c_c and

c_d of the sent symbols are not known to the decoder in the receiver and they must be evaluated in the decoding process. Calculated values of erasures Z_a, Z_b, Z_c and Z_d which will be used to correct received symbols at erasure positions must fulfill the following conditions

$$\begin{aligned} c_a &= v_a + Z_a\,, \\ c_b &= v_b + Z_b\,, \\ c_c &= v_c + Z_c\,, \\ c_d &= v_d + Z_d\,. \end{aligned} \tag{10}$$

Syndromes are evaluated based on the following set of equations

$$S_k = \sum_{i=0}^{q-2} \alpha^{ki} \hat{c}_i + \hat{p}_k\,, \quad k \in (0,4) \tag{11}$$

After calculating syndromes, the next decoding step is to form a set of syndrome equations in order to calculate the correction values for erasure corrections. We will suppose that the encoder and decoder agreed on a protocol in advance. Therefore, the decoder knows the control matrix (2), which could be expressed in a compact way as

$$\boldsymbol{H} = \left[\boldsymbol{H_P} \,\vdots\, \boldsymbol{I} \right] \tag{12}$$

where $\boldsymbol{H}_p$ is the parity part of the control matrix (2) and $\boldsymbol{I}$ is the identity 5×5 matrix. The detector supplies the decoder with the number of erasures ζ and their respective positions in the received vector $\boldsymbol{v}$. The erasure correcting algorithm then proceeds as follows:

1. If $\zeta = 0 \rightarrow$ (end of decoding), there are no erasures in the received vector, therefore it can be delivered as decoded or as an estimated codeword, else $\rightarrow$ go to step 2
2. If $1 \leq \zeta \leq 4 \rightarrow$ go to step 3, else $\rightarrow$ end of decoding (decoding failure – the code distance does not allow us to correct more than 4 erasures)
3. Evaluation of syndromes S_0, S_1, S_2, S_3 using (11). If $S_0 = S_1 = S_2 = S_3 = 0 \rightarrow$ end of decoding (it indicates decoding failure - there is a discrepancy between delivered message: $1 \leq \zeta \leq 4$ and calculated syndrome values), else $\rightarrow$ go to step 4
4. Out of matrix (2) create "erasure" matrix $\boldsymbol{H}_z$ so that its columns are columns of (2) corresponding to the respective erasure positions in $\boldsymbol{v}$. (There is a one to one correspondence between its rows and syndromes). $dim\{\boldsymbol{H}_z\} = 5 \times \zeta$, and $\rightarrow$ go to step 5
5. Find a $\zeta \times \zeta$ submatrix denoted as $\boldsymbol{H}_\zeta$ of $\boldsymbol{H}_z$ with nonzero determinant, and $\rightarrow$ go to step 6
6. Solve the system of linear equations: $\boldsymbol{S} = \boldsymbol{Z} \times \boldsymbol{H}_\zeta^\top$ for $\boldsymbol{Z}$, where $\boldsymbol{S}$ is a syndrome vector containing syndromes from $\{S_0, S_1, S_2, S_3\}$ corresponding to rows of $\boldsymbol{H}_z$ contained in $\boldsymbol{H}_\zeta$, $\boldsymbol{Z}$ is a vector of erasures, $\boldsymbol{H}_\zeta^\top$ is a transposed matrix $\boldsymbol{H}_\zeta$, and $\rightarrow$ go to step 7
7. By using calculated values of $\boldsymbol{Z}$ (obtained in the previous step) and (10) correct ocurred erasures, and $\rightarrow$ end of decoding

4 Conclusion

In this paper a decoding algorithm was presented for the recently discovered codes described in [5]. This algorithm allows correcting up to 4 erasures in each codeword of these codes.

Acknowledgments

This work was supported by the Scientific Grant Agency of Ministry of Education of Slovak Republic and Slovak Academy of Sciences (grant VEGA 1/0477/18).

References

[1] I. S. Reed and G. Solomon, "Polynomial Codes over Certain Finite Fields", *Journal of the Society for Industrial and Applied Mathematics* vol. 8, no. 2, pp. 300-304, 1960.

[2] S. B. Wicker, "Reed-Solomon Codes and their Applications", *IEEE Press*, Piscataway, NJ, USA, 1994.

[3] Kees A. Schouhamer Immink, "Codes for Mass Data Storage Systems", *Shannon Foundation Publisher*, 2004.

[4] W. Huffman and V.Pless, "BCH and ReedSolomon codes", *Fundamentals of Error-Correcting Codes*, pp. 168–208, Cambridge: Cambridge University Press, 2003.

[5] M. Rakús, P. Farkaš, T. Páleník and A. Daniš, "Five Times Extended Reed-Solomon Codes Applicable in Memory Storage Systems", *IEEE Letters of the Computer Society*, 2019.

[6] R. S. Elagooz, A. Mahran, S. Gasser and M. Aboul-Dahab, "Efficient Low-Complexity Decoding of CCSDS ReedSolomon Codes Based on Justesens Concatenation", *IEEE Access* vol. 7, pp. 49596–49603, 2019.

[7] A. Belyaev and P. Poperechny, "Reed-Solomon Encoder Design by Means of the Digital Filtration", *2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering* (EIConRus), Saint Petersburg and Moscow, Russia, pp. 2204–2207, 2019.

[8] M. Rajab, S. Shavgulidze and J. Freudenberger, "Soft-Input Bit-Flipping Decoding of Generalised Concatenated Codes for Application in Non-Volatile Flash Memories", *IET Communications* vol. 13, no. 4, pp. 460–467, 5 3 2019.

[9] B. Wang, P. Chen, Y. Fang and F. C. M. Lau, "The Design of Vertical RS-CRC and LDPC Code for Ship-Based Satellite Communications On-the-Move", *IEEE Access* vol. 7, pp. 44977–44986, 2019.

[10] G. Luo, X. Cao and X. Chen, "MDS Codes With Hulls of Arbitrary Dimensions and Their Quantum Error Correction", *IEEE Transactions on Information Theory* vol. 65, no. 5, pp. 2944–2952, May 2019.

[11] W. Halbawi, Z. Liu, I. M. Duursma, H. Dau and B. Hassibi, "Sparse and Balanced ReedSolomon and TamoBarg Codes", *IEEE Transactions on Information Theory* vol. 65, no. 1, pp. 118–130, Jan 2019.

[12] S. Ramabadran, A. S. MadhuKumar, G. Wang and S. K. Ting, "Joint Reconstruction of Reed-Solomon Encoder and Convolutional Interleaver in a Noisy Environment", *2018 International Symposium on Information Theory and Its Applications* (ISITA), Singapore, pp. 683–687, 2018.

Received 13 April 2019