ISSN 1064-2269, Journal of Communications Technology and Electronics, 2010, Vol. 55, No. 3, pp. 302–306. © Pleiades Publishing, Inc., 2010.
Original Russian Text © V.V. Zolotarev, G.V. Ovechkin, 2010, published in Radiotekhnika i Electronika, 2010, Vol. 55, No. 3, pp. 324–329.

THEORY AND METHODS OF SIFNAL PROCESSING

Efficient Multithreshold Decoding of Nonbinary Codes

V. V. Zolotarev and G. V. Ovechkin

Received April 29, 2009

Abstract—The main principles of multithreshold decoding (MTD) are generalized to nonbinary codes. The lower bounds of the probability of erroneous decoding are proposed. The nonbinary MTD efficiency is shown to be close to the values attained via optimum search methods that cannot usually be implemented for nonbinary codes. The complexity of the implementation of the proposed decoders is discussed.

DOI: 10.1134/S1064226910030083

INTRODUCTION

It has been shown [1–6] that, in many cases, multithreshold decoders (MTDs) and optimum decoding (OD) have almost identical characteristics in Gaussian channels with binary phase keying, exhibiting nearly coincident values even at a high noise level. However, in real systems, it is often convenient to operate with data having a byte structure. For example, byte data are used in storage systems with great amounts of information (archives recorded on optical disks or other carriers). In such systems, it is reasonable to apply nonbinary noise-immune codes for data protection against errors.

Until very recently, there were no efficient and relatively simple encoders and decoders of nonbinary (symbolic) data, except for methods based on Reed–Solomon (RS) codes. However, short RS codes with length n up to 255 symbols do not ensure the reliability that are now required. Decoders of long RS codes are extremely complicated, and their substantial simplification is barely feasible. Over the last few years, many researchers have actively developed decoders of nonbinary low-density codes [7–9]. These methods offer high correcting abilities. However, the increased complexity of implementation, especially when alphabet size q is large, impedes their application in real systems.

In this paper, MTDs are generalized to nonbinary symmetric channels [6, 10, 11]. At very high levels of noise, a nonbinary MTD (below, designated as qMTD) algorithm is demonstrated to be capable of ensuring such error probabilities of decoding that cannot be reached with arbitrarily long RS codes. In addition, it is found that the computational complexity of this algorithm is rather low and increases linearly with code length, i.e., will not exceed the theoretically determined minimum value.

1. BASIC THEOREM FOR THE NONBINARY MULTITHRESHOLD DECODER

Let us pass to a formalized description of the multithreshold algorithm for decoding of nonbinary codes [3, 4].

A q-ary ($q > 2$) symmetric channel (qSC) is assumed to be specified with the error probability $P_0 > 0$ such that any initial transmitted symbol of a codeword is changed randomly, independently, and equiprobably into one of the $q - 1$ remaining symbols. When any message is transmitted over the qSC, an optimum decision is, presumably, the single codeword belonging to q^{nR} possible codewords (n is the code length and R is the code rate, $R < 1$), which differs from the received message in the minimum number of code symbols.

Let us consider the nonlinear nonbinary systematic code whose $(n-k) \times n$ check matrix $\mathbf{H}$ is represented as

$$\mathbf{H} = [\mathbf{C} : -I_{n-k}],$$

Here, k is the length of the information part of the codeword; I_{n-k} is the $(n-k) \times (n-k)$ unit matrix; $\mathbf{C}$ is the $(n-k) \times k$ matrix containing zeros and ones, the ith row of which determines information symbols; and operation $[\mathbf{A} : \mathbf{B}]$ determines the matrix obtained via the horizontal concatenation of matrices $\mathbf{A}$ and $\mathbf{B}$. Let matrix $\mathbf{H}$ correspond to the binary self-orthogonal code whose matrix $\mathbf{C}$ is thoroughly described in [4]. Since check matrix $\mathbf{H}$ (and, consequently, the generating matrix $\mathbf{G} = [\mathbf{I}_k : \mathbf{C}^T]$) contains only zeros, ones, and negative ones, encoder and decoder operations applied to construct checking code symbols and to calculate $(n-k)$-long syndrome vector $\mathbf{S}$ of a received message are only summation and subtraction operations. Thus, encoding and decoding processes do not require the presence of a nonbinary field. It suffices to create any variant of a group of integers. For example, all the summation and subtraction operations can be performed with integers from the certain group modulo q, thereby substantially simplifying all the encoding procedures and the implementation of subsequent decoding.

Let the vector $\mathbf{Q} = \mathbf{A} + \mathbf{E}$, where $\mathbf{E}$ is the error vector, comes to a decoder after n-long code vector $\mathbf{A}$ with k information symbols is transmitted over the qSC. As in the binary case, each column vector $\mathbf{X}$ of length n can be represented by a pair of vectors $\mathbf{X}_I$ and $\mathbf{X}_V$ with lengths k and $(n-k)$, respectively. Thus, $\mathbf{X} = [\mathbf{X}_I; \mathbf{X}_V]$. Here, operation $[\mathbf{A}; \mathbf{B}]$ determines the matrix obtained via the vertical concatenation of matrices $\mathbf{A}$ and $\mathbf{B}$.

It is assumed that $\mathbf{D}$ is defined as the q-ary unit vector of length k:

$$\mathbf{D} = \mathbf{A}_I - \mathbf{Q}_I,$$

where $\mathbf{A}_I$ is the information part of the transmitted codeword $\mathbf{A} = [\mathbf{A}_I; \mathbf{A}_V]$ and $\mathbf{Q}_I$ is the information part of the received message $\mathbf{Q} = [\mathbf{Q}_I; \mathbf{Q}_V]$. Then, the following lemma is valid.

Lemma.

$$[\mathbf{D}; \mathbf{H} \times [\mathbf{Q}_I + \mathbf{D}; \mathbf{Q}_V]] = \mathbf{A} - \mathbf{Q}. \quad (1)$$

Proof. The code linearity implies that the equalities

$$\mathbf{S} = \mathbf{H} \times [\mathbf{Q}_I + \mathbf{D}; \mathbf{Q}_V] = \mathbf{H} \times [\mathbf{Q}_I + \mathbf{D}; \mathbf{Q}_V + \mathbf{A}_V - \mathbf{A}_V] = \mathbf{H} \times \mathbf{A} + \mathbf{H} \times [\mathbf{0}_I; \mathbf{Q}_V - \mathbf{A}_V],$$

where $\mathbf{0}_I$ is the zero vector of length k, are valid.

For a systematic code,

$$\mathbf{H} \times [\mathbf{0}_I; \mathbf{X}_V] = -\mathbf{X}_V,$$

From this condition, it can be inferred that $\mathbf{S} = \mathbf{A}_V - \mathbf{Q}_V$. Since $\mathbf{D} = \mathbf{A}_I - \mathbf{Q}_I$, we obtain $[\mathbf{D}; \mathbf{S}] = \mathbf{A} - \mathbf{Q}$. The lemma is proved.

By analogy with the binary case, this lemma provides a simple and useful relationship between an arbitrary codeword and the received message and, consequently, states that, for codewords under consideration, syndrome vector $\mathbf{S}$ is the check-symbol difference between received message $\mathbf{Q}$ and code vector $\mathbf{A}$ with information part $\mathbf{A}_I$. Such an interpretation of the syndrome vector was estimated from different standpoints in [1, 4, 5].

The lemma proved above enables us to prove the main property for the qMTD algorithm described below.

Let a decoder receive the vector $\mathbf{Q} = \mathbf{A} + \mathbf{E}$ that involves distortions of codeword $\mathbf{A}$ caused by its transmission over the qSC. Similar to the binary case, q-ary difference vector $\mathbf{D}$ is taken to be $\mathbf{0}_I$ at the beginning of a decoding procedure. After calculating the syndrome vector $\mathbf{S} = \mathbf{H} \times \mathbf{Q}$ of a received message by means of usual methods, the qMTD is assumed to perform processing according to the following scheme.

(i) For arbitrary chosen q-ary decoded information symbol i_j, the decoder estimates the number of two most frequently observed values of checks in the total number J of checks associated with symbol i_j and the symbol d_j of vector $\mathbf{D}$ corresponding to symbol i_j. It is assumed that the values of these two checks are h_0 and h_1 ($0 \le h_0$ and $h_1 \le q$) and they are repeated, respectively, m_0 times and m_1 times ($m_0 \ge m_1$). This procedure is similar to the calculation of the sum of checks in the thresholder of a binary MTD.

(ii) If $m_0 - m_1 \le T$, where $T = 0, 1, 2, \ldots$ is a nonnegative integer, any new symbol i_m ($m \ne j$) is selected and processed, starting with step (i). This procedure is similar to comparison with the threshold in the binary decoder.

(iii) If $m_0 - m_1 > T$, error estimate h_0 is subtracted from symbols i_j and d_j and from total number J of checks corresponding to i_j. Then, new symbol i_m, $m \ne j$is selected and the transition to step (i) is performed.

At this final step of a decoding cycle, the symbol decoded is changed and all the syndrome symbols—the checks with respect to the symbol decoded—are corrected through feedback. However, in contrast to a binary MTD, the summation and subtraction of a qMTD are nonidentity operations.

For each symbol of a received message, such attempts of decoding can be repeated three, ten, and more times.

As in the binary case, the qMTD algorithm can easily be implemented by successively estimating all the information symbols. The decoding procedure is stopped if the fixed number of error corrections is attained or any attempt of correction does not lead to a change in the value of the symbol decoded.

The qMTD algorithm discussed above satisfies the following theorem.

Main theorem of multithreshold decoding of nonbinary codes.

Let the nonbinary self-orthogonal code described above is processed by the decoder using the qMTD algorithm. Then, each change of symbols decoded leads to a more plausible decision when compared to the preceding state of the decoder.

Preliminary discussion. From two code vectors of a classical qSC, the code vector with the least number of symbols differing from those of the received vector can be regarded as more closely coincident with the received message and, consequently, more plausible. Hence, a growth in the plausibility of qMTD decisions with each change of symbols decoded can be proved by the fact that each new codeword has the increased number of symbols coinciding with the symbols of the received message. In other words, it is necessary to demonstrate that the Hamming distance between the codeword and the message decreases. For nonbinary symbols, this distance immediately corresponds to the number of noncoincident symbols in two vectors with equal lengths.

Thus, with allowance for the properties of a syndrome vector and difference register and according to the lemma for the qMTD, the Hamming distance between the received vector and the current decision of the qMTD is equal to the number of nonzero symbols of the syndrome vector and difference register. Hence, the deceased Hamming distance—the enhanced plausibility of decisions of this decoder—

can be attained by selecting another codeword that ensures the increased total number of zero symbols of syndrome **S** and difference vector **D**. It should be recalled that, similar to the binary case, information symbols of the codeword are assumed to reside in the corresponding registers of the decoder.

Proof. Let the decoder processes the vectors $\mathbf{A}_{1I}$, $\mathbf{D}_1 = \mathbf{A}_{1I} - \mathbf{Q}_I$, and $\mathbf{S}_1 = \mathbf{H} \times [\mathbf{Q}_I + \mathbf{D}_1; \mathbf{Q}_{\bar{I}}]$,

where $\mathbf{A}_1 = [\mathbf{A}_{1I}; \mathbf{A}_{1\bar{I}}]$ is an arbitrary codeword and **Q** is the received message.

Let us demonstrate that the qMTD algorithm changes any symbol i_j in the current information decision vector $\mathbf{A}_{1I}$ of the decoder, thereby generating new decision vector $\mathbf{A}_{2I}$. In this case, the Hamming distance between codeword $\mathbf{A}_{2I}$ and received vector **Q** is less than the same distance corresponding to decision vector $\mathbf{A}_1$; i.e., $|\mathbf{A}_1 - \mathbf{Q}| > |\mathbf{A}_2 - \mathbf{Q}|$.

Indeed, a change in some symbol i_j indicates that single check value h_0 ($h_0 \neq$) is strictly more frequently repeated (m_0 times) than all other values (m_1 times) in the set of checks of i_j ($m_0 > m_1$). In this case, after quantity h_0 is subtracted from i_j, d_j, and total number J of checks in the syndrome register, number m_0 of checks (and, presumably, symbol d_j) equal to h_0 vanishes. In the syndrome vector, the number of zero checks (undoubtedly, with the value of symbol d_j taken into account), which were equal to zero before the change of symbol i_j, cannot exceed m_1. This implies that, for these positions, the weight of the syndrome vector can increase by no more than m_1 when the change of i_j is caused by these symbols. Then, a total change in weight is $m_1 - m_0 < 0$; i.e., the total weight of vectors **S** and **D** decreases after a change in symbol i_j decoded. Note that vector $\mathbf{S}_2$ differs from vector $\mathbf{S}_1$ only in the symbols corresponding to checks of i_j and difference vectors $\mathbf{D}_2$ and $\mathbf{D}_1$, as well as the checks of symbol d_j, differ by quantity h_0 only in position i_j. However, this indicates that the qMTD contains vectors $\mathbf{S}_2$ and $\mathbf{D}_2$ corresponding to the difference between the received vector and the new solution of the decoder after the change of i_j; i.e., condition (1) is valid. Then, it is evident that, in the new state of the decoder, the conditions of the lemma are satisfied, making it possible to pass to the next attempt of correcting another symbol i_j ($m \neq j$). As a result of this attempt, the change of the symbol decoded again guarantees the transition to the more plausible decision, and so on. The theorem is proved.

Note two most substantial aspects of the proposed algorithm. The first aspect consists in that, similar to the case of binary codes, each of the decoding iterations will not obligatorily improve a decision until the OD decision is attained. Indeed, some error configurations of self-orthogonal codes can be corrected only via OD, rather than by the qMTD. Hence, the basic way to increase the qMTD efficiency is search for codes in which such uncorrectable error configurations are rarely observed even at a high noise level.

Another important aspect is that the presence of an absolutely strict majority of checks is not required. In comparison with a traditional approach to binary majority schemes, the symbol decoded by qMTD can be changed with the use of a relatively strict majority of checks, as follows from the condition $m_0 - m_1 > T$. For example, in the self-orthogonal codes with $d = 9$, the error of a symbol decoded can be corrected even if its nine checks (including symbol d_j of the difference register) contain only two true checks, while other seven are erroneous. Such correction cannot be implemented for binary codes, but is typical for the qMTD. In this example, the sole condition is the different values of checks when symbol i_j is decoded. At almost all large value of alphabet q, this condition is satisfied. This property of the qMTD substantially expands its abilities at high noise levels, simultaneously ensuring the fairly low complexity of majority procedures in q-ary channels.

2. LOWER BOUNDS OF THE ERROR PROBABILITY OF DECODING

Let us calculate the lower bound of the probability of optimum decoding when codes are specified in the form described above. For this purpose, it is necessary to reveal the most frequent events at which the Hamming distance between an error vector and the nearest nonzero codeword is less than its own weight. In the presence of a linear code, this condition is sufficient to make an erroneous decision even with the use of an optimum decoder. When an error vector exhibits such properties, it is necessary to analyze the code symbols corresponding to the positions of checks associated with current decoded symbol i_k.

Let us describe the probabilities of certain most simple events leading to OD errors [34, 6, 11]:

(i) If all the check symbols and decoded symbol i_k are erroneous,

$$P_1(e) = P_0^{J+1}, \tag{2}$$

where $d = J + 1$, d is the minimum code distance of the self-orthogonal code).

(ii) If all the check symbols are erroneous, but two of them are identical, and i_k is correct,

$$P_2(e) = \frac{J(J-1)(1-P_0)P_0^{J-2}}{2(q-1)} \prod_{i=1}^{J} \left(1 - \frac{i}{q-1}\right); \tag{3}$$

(iii) If one check symbol is correct received, but other symbols and i_k are erroneous,

$$P_3(e) = J(1-P_0)P_0^J. \tag{4}$$

As a result, the lower bound of the error probability for an optimum decoder is determined from the expression

$$P_{\text{opt}} = P_1(e) + P_2(e) + P_3(e). \tag{5}$$

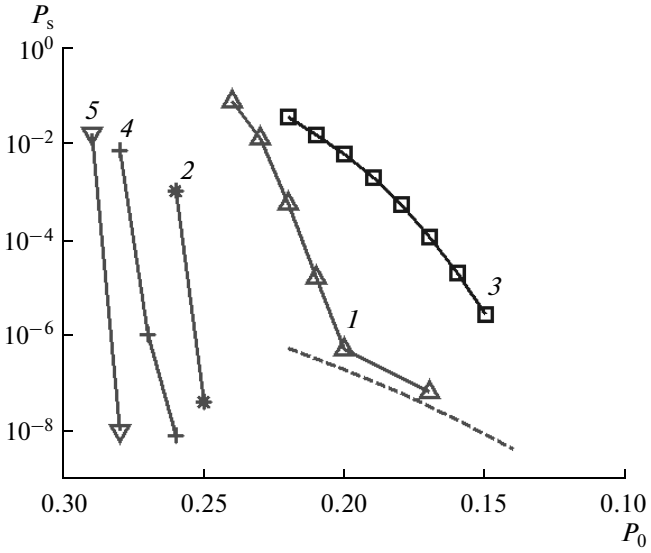
Different events leading to the errors of nonbinary OD, estimation of the probability of their occurrence, and the probabilities of the erroneous first symbol of a nonbinary threshold decoder, which can be regarded as the upper bounds of the error probability of the qMTD, have been analyzed more comprehensively in [3, 4, 6, 11].

In almost all the practical applications of codes, the number of events mentioned above is sufficient to obtain the fairly accurate probabilistic estimates of their potential noise immunity. Since each iteration step of the qMTD approximates to an OD decision, it may be expected that the required optimum decision will be attained at a certain very high level of noise. The search for the optimum decision necessitates such a number of examinations that increases exponentially with code length n. In the case of the qMTD, the decoder complexity remains the linear function of n, i.e., reaches the theoretically minimum level.

3. qMTD AND qSC CHARACTERISTICS

The dependences between the symbol error probability of the qMTD (P_s) and the symbol error probability in the q (P_0), which were obtained via computer simulation for codes with rate $R = 1/2$, are presented in the figure. Here, curves *1* and *2* correspond to the qMTD characteristics for codes whose lengths are 4000 and 32 000 one-byte symbols ($q = 256$) and code distances are, respectively, $d = 13$ and 17. During the error correction process, the number of decoding iterations varies from five to fifteen. In the figure, the dashed line corresponds to the lower bound of the symbol error probability of OD (P_{opt}) for the first code with $d = 13$. It is seen that the efficiencies of the qMTD and OD are closely coincident even when the noise level in the channel is very high. For comparison, the characteristics of the (255, 128) RS code with the same alphabet size $q = 256$ (curve *3*) are also presented in the figure. Note that, for the symbols of the same size, the efficiency of the qMTD is much higher than the efficiency of the RS code owing to the larger code lengths and good convergence of qMTD decisions to the OD decision. To attain these results with the use of the qMTD, decoding parameters and codes must be selected carefully. In the selection of codes, the main criterion is their resistance to the error-propagation effect, which manifests itself in the fact that the first error gives rise to a significant increase in the probability of subsequent errors. As a result, packet errors can appear at the decoder output. A detailed description of principles used to select codes resistant to the error-propagation effect can be found in [4].

The qMTD is superior to other error correction methods in that this decoder makes it possible to operate with symbols of an arbitrary large size and to ensure the same high correcting ability. This is confirmed by the qMTD characteristics presented in the figure, which were obtained for the code with $n = 32000$, $d = 17$, and two-byte (curve *4*) and four-byte (curve *5*) symbols.



qMTD characteristics for codes with $R = 1/2$ in the qSC. The behavior of curves *1*–*5* is described in Section 3.

Further significant improvement of the decoding efficiency of qMTD algorithm can be implemented by using convolutional codes, the serial and parallel concatenation encoding methods, codes with allocated branches, and other methods that are partially described in [3–5].

4. COMPLEXITY OF THE IMPLEMENTATION OF THE qMTD

Analysis of the nonbinary qMTD has confirmed the linearity decoding complexity. It is evident that the number of operations required to decode one symbol is determined only by the code distance, remaining independent of the code length and the size of used symbols. In the software implementation of the qMTD, its threshold elements is controlled by the subroutine that occupies less than ten short lines in the C^{++} language and ensures simultaneous processing of the bytes of a received or stored massage whose number depends on the architecture of the processor unit applied. The demoprogram for a symbolic qMTD is available on web site [3]. When operated on usual personal computer, the demoprogram demonstrates almost optimum decoding of very long codes at the data rate of more than 10 Mb/s. In addition, the demoprogram simulates all the operations performed in the data transmission channel: information stream formation, encoding, noise distortions, and operation of the decoding algorithm under consideration. Thus, the real processing power of the software-based qMTD can be regarded to be more than 1.5–2 times as great.

It should be emphasized that the complexity of the widespread decoders of RS codes is proportional to

the square of the code length. In addition, the efficiency of these codes is substantially less than the efficiency of the qMTD. The complexity of various methods used to improve the correcting ability of RS code, including all the variants of the Sudan algorithm, is on the order of n^3. For codes with a length of 30000 symbols, the difference in the complexity order is $n^2 = 30000^2 \approx 10^9$, i.e., about billion times as great. In this case, an increase in noise immunity is insignificant. Thus, it is shown that are all new and complicated methods used to decode RS codes are inefficient in comparison with the proposed qMTD.

CONCLUSIONS

Although Reed–Solomon codes were deservedly regarded as the best nonbinary codes for almost fifty years, the abilities of symbolic qMTD algorithms in error probability correction and the number of decoding operations exceeds those of RS codes by many orders of magnitude. This is caused by the efficient application of nonbinary multithreshold decoding ideas to very simply organized nonbinary codes of an arbitrary large length. As a result, the earlier unattainable level of noise immunity provided by different-type qMTD algorithms makes it possible to ensure the high reliability of data transmission and storage and to solve complicated problems without any supplementary modification of these algorithms, performing only insignificant adaptations to the requirements arising in large-scale digital systems.

ACKNOWLEDGMENTS

This study was supported the Russian Foundation for Basic Research, project no. 08-07-00078.

REFERENCES

1. S. I. Samoilenko, A. A. Davydov, V. V. Zolotarev, and E. I. Tret'yakova, *Computer Networks* (Nauka, Moscow, 1981) [in Russian].
2. V. V. Zolotarev and G. V. Ovechkin, Elektrosvyaz, No. 9, 34 (2003).
3. www.mtdbest.iki.rssi.ru
4. V. V. Zolotarev, *Theory and Algorithms of Manythreshold Decoding* (Radio i svyaz', Goryachaya liniya-Telekom, Moscow, 2006) [in Russian].
5. V. V. Zolotarev and G. V. Ovechkin, *Noiseless Encoding. Methods and Algorithms (Handbook))* (Goryachaya liniya-Telekom, Moscow, 2004) [in Russian].
6. V. V. Zolotarev, Mobil, Sist., No. 3, 25 (2006).
7. A. Bennatan and D. Burshtein, IEEE Trans. Inf. Theory **52**, 549 (2006).
8. F. Zhang and H. D. Pflster, in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM'07), Washington, DC, Nov. 26–30, 2007* (IEEE, New York, 2007), p. 283.
9. D. Declercq and M. Fossorier, in *Proc. Int. Symp. Information Theory (ISIT 2005), Adelaide, Sep. 4–9, 2005* (IEEE, New York, 2005), p. 464.
10. www.mtdbest.iki.rssi.ru/pdf/qmtd_iscta07.pdf
11. V. V. Zolotarev, Mobil. Sist., No. 3, 39 (2007).
12. R. L. Townsend and E. J. Weldon, IEEE Trans. Inf. Theory **13**, 183 (1967).
13. M. Sudan, J. Complexity **13**, 180 (1997).
14. V. V. Zolotarev, Mobil. Sist., No. 3, 66 (2008).