A New Method on Channel Coding Category Identification

Dong Zhang[1, a], Zhengwei Lei[1], Yanmei Lv[1], Lijuan Peng[1]

[1]Machine Technology Institute, Shijiazhuang, China

[a]310792132@qq.com

Keywords: channel coding; detection and identification; analyzing matrix

Abstract. The channel coding is a key link in modern digital communication systems, which effectively improves the reliability of communication. In view of the problem of the categories detection of channel coding, this paper puts forward a method of detecting channel coding categories using the correlation characteristics of coding sequences. The method constructs the correlation characteristics matrix for coding sequence, and judges the channel coding categories of sequence through the simplified results of the matrix. The simulation result shows that this method can effectively detect channel coding categories while requires a small amount of data and simple calculation.

Introduction

Currently, the widely used channel coding mainly includes Hamming codes, Cyclic codes, BCH codes, RS codes, LDPC codes, Convolution codes and Turbo codes. While the widely using of channel coding, the identification of code categories has become a problem to solve in these fields such as information recovery, individual characteristics analysis and information confrontation. At present, many domestic and foreign scholars carried on the thorough research to the problem of the channel coding identification. Literature [1] puts forward the method using dual code recognition for the (n, K) linear block codes identification problem, which has certain fault tolerance. Literature [2] establishes the identification model for cyclic code and proposes an iterative identification method under the condition that the code length is known. The literature [3,4] respectively study the identification problems of the two special linear block codes, BCH codes and RS codes. Literature[5，6] discuss the identification problem of Turbo code, which have certain guidance. Literature[7] presents the channel coding detection algorithm, which can distinguish whether the received sequence are coded.

The current related research focuses on parameter identification of a specific encoded method, which are mostly based on the known channel coding categories. So the methods can't be used when the channel coding category is unknown. This paper proposes a channel coding category identification method which can effectively identify the channel coding type of the known sequence, which lays a foundation for the further identification of the specific parameters.

The Principle and Characteristics of Common Channel Coding

Channel coding in practical applications can be divided into two categories, linear block codes and convolution codes. The former based on finite field algebra theory, with strict mathematical expression, and can be divided into Hamming code, cyclic code, BCH code, RS code and LDPC code and so on according to encoding modes [8]. The latter introduces check relationship among coding sequences by convolution operation, which has not complete theory foundation as the former.

Commonly used linear block codes. The linearity in the Linear block codes means the constraint relation among code elements is linear, and the Block means that in coding every k information bits were independently processed as a group, which are transformed into a group of binary code of length n. So the parameters of linear block codes is generally expressed as (n, K), where n is the length of code group, and K is length of information, coding process as Eq.1.[8]

$$c = mG \tag{1}$$

In Eq.1 C is a linear block codes, m represents any k-dimensional information vector, and G is the generator matrix with k lines and n column.

$$G = \begin{bmatrix} g_{1,1} & \cdots & g_{1,n} \\ \vdots & \ddots & \vdots \\ g_{k,1} & \cdots & g_{k,n} \end{bmatrix} \tag{2}$$

The following analysis on different linear block codes bases on the Eq.2 and Eq.3. There is specific constraint relationship between the length of information bits k and the length of code words n for Hamming code. That is, $n = 2^m - 1$, $k = 2^m - m - 1$, in which m is an integer greater than 2. Hamming code belongs to the complete code, whose generator matrix can be obtained by equivalent transforming of the following system generator matrix.

$$G = \begin{bmatrix} 1 & 0 & \cdots & 0 & g_{1,k+1} & \cdots & g_{1,n} \\ 0 & 1 & \cdots & 0 & g_{2,k+1} & \cdots & g_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & g_{k,k+1} & \cdots & g_{k,n} \end{bmatrix}_{(2^m-1-m)\times(2^m-1)} \tag{3}$$

Different from the Hamming code, the values has no constraints of the parameters k and n for Cyclic code, and its generator matrix like Eq.4.

$$G = \begin{bmatrix} g_0 & g_1 & g_2 & \cdots & g_{n-k} & 0 & 0 & \cdots & 0 \\ 0 & g_0 & g_1 & \cdots & g_{n-k-1} & g_{n-k} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & g_0 & g_1 & \cdots & g_{n-k} \end{bmatrix}_{k \times n} \tag{4}$$

Among them, cg_i is the coefficient of the generator polynomial $g(x) = g_o + g_1 x + \cdots + g_{n-k} x^{n-k}$ of the Cyclic codes. Generally, $cg_o = g_{n-k} = 1$, the rest of the coefficients are 0 or 1. In addition, it is worth noting that the Hamming code belongs to Cyclic code whose length of information bits and length of code has a special relationship.

Convolution code. Convolution encoder outputs n code elements in any given unit of time. And each element not only associated with the k information codes the encoder inputs in the time unit, but also associated with the information inputted in the consecutive m time units before. So Convolution code is generally represented by three parameters: (n,k,m). This paper only studies the Convolution code in binary domain $GF(2)$.

Convolution code can be constructed using the generator Polynomial matrix. At first the information sequence I is expressed as a polynomial $I(x) = [I_1(x), I_2(x), \cdots, I_k(x)]$, then the coding sequence c represented by polynomial as $c(x) = [c_1(x), c_2(x), \cdots, c_n(x)]$, we can establish Eq.5.

$$c(x) = I(x)G(x) \tag{5}$$

From the analysis above we can find that the definition of code length in Convolution code is different from the linear block codes. In the practical application the value is not greater than 8[9].

Detection and Identification of Channel Coding Categories

This section presents a channel coding categories detection and identification method based on the analysis of the theory and the characteristics of several common channel coding. At first estimates the

code length by the known sequence, then distinguish convolution code or linear block code, and then identify which kind of channel coding according their respective characteristics.

The recognition of the channel coding needs the check matrix H corresponding to the generator matrix. The H is $(n-k) \times n$ dimension matrix for the linear block code, and which is polynomial matrix for Convolution code. So we know that the relationship between the check matrix and the code word sequences is Eq.6.

$$cH^{\mathrm{T}} = 0 \tag{6}$$

Construct matrix C ($N \times (2n+2)$ dimension and $N > 2n+2$) with the known coding sequence.

The Eq.9 shows that when the code length is n, $CH^{\mathrm{T}} = 0$, that is , the homogeneous equation exists a non-zero solution when the matrix C is the coefficients. So that we can determine whether n is the code length by solving the rank of the matrix C. It should be noted that, the number of columns of Convolution code matrix C should be bound length, rather than the code length n in the definition because the convolution code is semi-infinite length coding sequence.

Based on the discussion above, the detection and identification method of channel coding categories is shown in Fig. 1, the steps are as follows:

1) Assuming the code length is n, whose value is set 2~8. Construct the $(12n+2) \times (12n+1)$ dimension code-word matrix C like Eq.7 according to the code length n.

$$C = \begin{bmatrix} c_1 & c_2 & \cdots & c_{12n+1} \\ c_{n+1} & c_{n+2} & \cdots & c_{13n+1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{(12n+1)n+1} & c_{(12n+1)n+2} & \cdots & c_{(12n+13)n+1} \end{bmatrix} \tag{7}$$

2) The Eq.8 shows the result of simplifying the matrix C with addition modulo 2 (only row operation), and check whether the rank K is less than 12n. If yes, then judge whether the code is convolution code or linear block codes. When n=7 and $(12 \cdot 3 - 3) \le K \le (12 \cdot 4 + 4)$, it is the linear block codes with code length 7 , or it is the convolution code. If K is not less than 12n and n is less than 8, return to the step 1, or execute step 3;

$$C' = \begin{bmatrix} I_v & 0 & 0 & 0 \\ 0 & I_k & P & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_w \end{bmatrix} \tag{8}$$

3) As the code length n > 8 , we can preliminary judge that the code is the linear block code. Reset the code length n by setting the values of 9~1023, and then construct the matrix C with dimension as $(n+2) \times (n+1)$ according to step 1.

4) Simplify the matrix C to C', and judge if it is the same as Eq.11. If yes and P's dimension is $k \times (n-k)$ $n = 2^m - 1$, $k = n - m$, then it is (n, k) Hamming code; else if P's dimension is $k \times (n-k)$ $n = 2^m - 1$, k≠ n-m , then the it is BCH code. And if $n \ne 2^m - 1$, then it is common Cyclic code. Set K equals to the rank of C', if $K \ge n$ and n＜1023, then return to step 3, or finish the process.

Simulation

This section introduces the validation of channel coding identification method by simulating. In the simulation channel coding selects commonly used parameters. Taking examples for Hamming codes, convolution codes, general circulation code and primitive BCH codes, the specific parameters are listed in Table 1.

Putting the random 0,1 sequences with length of 500 as an information sequence to be encoded, respectively performs the four channel coding using parameters in Table 1, and generates encoded sequence c1, c2, c3 and c4. Construct matrix C based on the method discussed in the third section. For code word sequence c1 generated by the first type channel coding method, when n is set 2~ 6, the simplification of matrix C is shown in Fig. 2 (with n = 3 as an example), when n = 7 simplified results is shown in Fig. 3. Comparing the two figures can clearly identify that the channel coding is (7, 4) Hamming code.
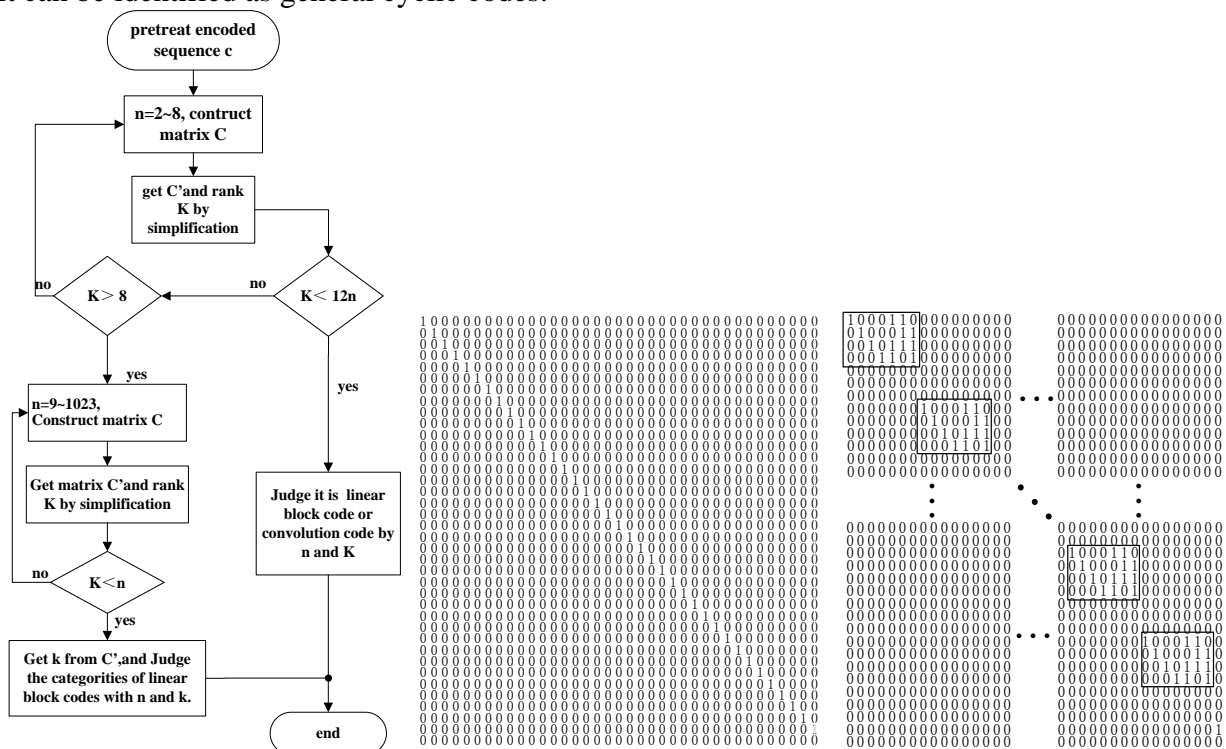
For the code sequences of the second channel coding type, simplification results of the constructed matrix is shown in Fig. 4 when n = 2. It shows that only the (2,1, m) convolution code check matrix is $2m \times 1$, it is able to determine that the coding mode is (2,1,7) convolution code.

When using the third channel coding type, the simplification results are shown in Fig. 5 only when $n = 47 \neq 2^m - 1$, other values of n corresponding simplification results are like which is show in Fig. 2. Fig. 5 shows that the channel coding is linear block code, and k, the length of information bits, is 24, so it can be identified as general cyclic codes.

Fig. 1 Channel coding detection   Fig. 2 simplification results   Fig. 3 simplification results

Table 1 channel coding parameters

channel coding	Parameters (n, k) or (n, k, m)
Hamming code	(7,4)
convolution code	(2,1,7)
cyclic code	(47, 24)
BCH code	(31,21)

If using the fourth channel coding type, the simplification result is shown in Fig. 6 only when $n = 31 = 2^5 - 1$, other values of n corresponding simplification results are similar to Fig. 2. The figure shows its information bits length is 21, it can be identified as BCH codes.

The simulation above shows that the proposed method can effectively distinguish the common several types of channel coding type, which laying the foundation for further estimation of the specific parameters.

Summary

This paper proposes the detection and identification method of channel coding categories based on the analysis of the theory and characteristics of several common channel coding. First construct analysis matrix for the code sequences according to the assumption code length, then judge whether the code length is true through simplifying the matrix. If the value is true, then get the length of information bits, and then determine channel coding categories. The simulations show that this method can effectively detect channel coding categories. And it requires a small amount of data and simple calculation when the code length n is small. Therefore, The channel coding method for detection and identification is of certain value, which lays the foundation for further parameter estimation.

Fig. 4 (2,1,7) convolution code when n=47

Fig 5 (47, 24) cyclic code when n=31

Fig 6 (31,21) BCH when n=2

References

[1] Valembois A. Detection and recognition of a binary linear code [J] Discrete Applied Mathematics, 2001, 111 (1):.199-218.

[2] Wang J, Yue Y, Yao J. A method of blind recognition of cyclic code generator polynomial [C] Wireless Communications Networking and Mobile Computing (WiCOM), 2010 6th International Conference on IEEE, 2010:.. 1-4.

[3] Zhang Yongguang, Zheng Shilian , study on parameters identification of BCH code [J] .Journal of Xidian University (Natural Science), 2013,40 (5): 72-77.

[4] Xie hui, Wang Fenghua, etc. RS code blind detection identification method based on the spectrum pretreatment [J] .Journal Acta Astronautics, 2013,34 (1): 128-132.

[5] Zhang Yongguang .One blind identification method of parameter coding of Turbo code [J] .Journal of Xidian University (Natural Science),2011, 38 (2): 167-172.

[6] Li Xiaotian, Zhang Runsheng, Li Yanbin . Identification algorithm of Turbo return to zero [J] .Journal of Xidian University (Natural Science),2013, 40 (4): 161-166.

[7] Moosavi R, Larsson E G. A Fast Scheme for Blind Identification of Channel Codes [C] Global Telecommunications Conference (GLOBECOM2011), 2011 IEEE IEEE, 2011:.1-5.

[8] Zhao Xiaoqun. Modern coding theory [M] Wuhan: Huazhong University of Science and Technology Press, 2008.

[9] CCSDS / 131.0-B-1-2003, CCSDS Recommendation for TM Synchronization and Channel Coding [S] Washington: CCSDS Secretariat,2003.