Generalized concatenated coding and Fourier transform

Valentine Afanassiev, Alexander Davydov, Vladimir Potapov
{afanv,adav,potapov}@iitp.ru
Institute for Information Transmission Problems (Kharkevich Institute), Russian Academy of Sciences, Bol'shoi Karetnyi per. 19, GSP-4, Moscow, 127994, RUSSIA

Dedicated to the memory of Professor Stefan Dodunekov

Abstract. A general platform for Generalized concatenated code (GCC) encoding and decoding based on multidimensional Fourier transform is considered. A description of product codes (PC) sets that are embedded into a GCC and vice-versa is given. The rules for transition from a layer of GCC to embedded PC and from a definite PC layer to embedded layers of GCC are defined. We demonstrate how it is possible to use the transition rules for advanced decoding of a GCC and how it is reasonable to combine an iterative decoding of embedded PC layers with a standard steps for GCC decoding.

1 Introduction

Let us start with well known definition of Reed-Solomon (RS) codes [1]. Let the data be a sequence of k symbols, interpreted as coefficients of a polynomial $U(x) = \sum_{i=0}^{k-1} U_i x^i$ of the degree $k-1$ (or less) over a finite Galois Field $GF(q)$, $n|(q-1)$. The transmitted codeword is then a sequence of $n > k$ values attained by this polynomial in n district points: $\boldsymbol{C} = [U(x_0), U(x_1), ..., U(x_{N-1})]$. Two distinct codewords can agree in at most $k-1$ points, since the difference polynomial can have at most $k-1$ roots. An equivalent definition is just inverse Fourier Transform of the vector $\boldsymbol{U} = (U_0, ..., U_{N-1})$.

Thus we have $\boldsymbol{C} = \boldsymbol{U}\boldsymbol{\Phi}^{-1}$, where $\boldsymbol{\Phi}$ is the Fourier Transform matrix. In the canonic form $\boldsymbol{\Phi} = \left(\alpha^{ij}\right)$, $i, j = 0, 1, \ldots, n-1$, where $\alpha^n = 1$ be an element of order n of a finite field, $\boldsymbol{\Phi} = (\boldsymbol{\Phi})^T$. The inverse Fourier transform is given with matrix $\boldsymbol{\Phi}^{-1} = \left(\alpha^{-ij}\right)$, $i, j = 0, 1, \ldots, n-1$. It is clear that $\mathbf{C\Phi} = \mathbf{U\Phi}^{-1}\Phi = \mathbf{U}$. It also evident that the given definition gives us the set of embedded RS codes for any $k = 1, 2, ..., n$.

Theory of Concatenated Codes opened by Forney in [5] was generalized later to theory of Generalized Concatenated Codes (GCC) in [2] and [3], see also [6]. One special case was described in [1] as two-dimensional Fourier Transform over $GF(q)$. Here we consider GCC as two-dimensional Fourier Transform over $GF(q)$ and demonstrate same special properties of the construction that are useful for advanced decoding. The first of them is symmetry (in row-column space) and the second one is great freedom in decomposition of the general GCC on a *serially embedded subcodes.*

Most of problems of optimal designing and decoding of GCC was formulated and analyzed in [4]. Here we going to open (explain) new possibilities without detailed comparison and complexity analysis.

2 The structure of Generalized Concatenated Code (GCC) based on Fourier Transform

A codeword of GCC over $GF(q)$ is defined as $n \times n$ matrix calculated as two-dimensional Fourier Transform over $GF(q)$ of information symbols matrix $\mathfrak{I} = (a_{i,j})$ bounded by k rows and k columns, $i, j = 1, ..., k < n - 1$,

$$\mathbf{C} = \mathbf{\Phi}^{-1}\mathfrak{I}\mathbf{\Phi}^{-1}$$

We use here a symmetric (square) form of $\mathfrak{I}$ just as a simplest example. The infilling configuration of $\mathfrak{I}$ can be different. If it is complete square then we get after two-dimensional Fourier Transform a Product Code (PC) of the length n^2 and dimension k^2 with the code distance $(n - k + 1)^2$. If it is upper triangular form then we get a GCC of the length n^2 and dimension $\left(k^2 - k\right)/2$ with the code distance $\max_i (n - i + 1)(k - i + 1)$. There are number of intermedium variants for $\mathfrak{I}$ configurations. Formally, all infilling configurations are available. At this point it is important that all possible configurations of $\mathfrak{I}$ give us different codes that all are embedded in the Product Code (PC) of the length n^2 and dimension k^2.

Let us define the parameter k as level of the two-dimensional code. So, $\mathbf{C}_k = \mathbf{\Phi}^{-1}\mathfrak{I}_k\mathbf{\Phi}^{-1}$. In a general case form of $\mathfrak{I}$ matrix can be any rectangular. In that case level of a code has be different on row and on column direction. Let we have a code $\mathbf{C}_k$ embedded in PC of the level k . Now we can expand the given code to GCC of the level $k + 1$ as follows. Define an $n \times n$ matrix $\Delta_k = (\delta_{ij})$ the only nonzero elements of which are $\delta_{ik}, i < k$, and/or $\delta_{k,j}, j < k$.

$$\mathbf{C}_{k+1} = \mathbf{C}_k + \mathbf{\Phi}^{-1}\mathbf{\Delta}_{k+1}\mathbf{\Phi}^{-1} = \mathbf{\Phi}^{-1}\left(\mathfrak{I} + \mathbf{\Delta}_{k+1}\right)\mathbf{\Phi}^{-1}$$

The last equation can be used recursively and it gives the way for decomposition of a given $\mathbf{C}_k$ on a sequence of embedded PC and/or GCC. In dependence on form of Δ_i, $i < k$, here can be a square or rectangular forms of a resulting codes. So, to make a transition from $\mathbf{C}_k$ to an embedded code $\mathbf{C}_{k-1}$ it is enough to find the $\mathbf{\Delta}_k$ matrix and substract its two-dimensional Fourier Transform. We call the matrix $\mathbf{\Phi}^{-1}\mathbf{\Delta}_k\mathbf{\Phi}^{-1}$ as k -th layer of GCC or PC.

Consider the step by step encoding of a GCC.

The first one (the first Fourier Transform) can be like this

$$\mathbf{A}_{Row} = \mathfrak{I}_k \mathbf{\Phi}^{-1} = \begin{bmatrix} \mathbf{A}_{k\times n} \\ \mathbf{0}_{n-k\times n} \end{bmatrix},$$

or

$$\mathbf{A}_{Col} = \mathbf{\Phi}^{-1}\mathfrak{I}_k = \left[\mathbf{A}_{n\times k}\mathbf{0}_{n\times n-k}\right].$$

The second one (the second Fourier Transform) must be like this:

$$\mathbf{C} = \left(\mathbf{\Phi}^{-1}\right)^T \mathfrak{I}_k \mathbf{\Phi}^{-1} = \mathbf{\Phi}^{-1}\mathbf{A}_{Row} = \mathbf{A}_{Col}\mathbf{\Phi}^{-1} \in \mathrm{PC}. \tag{1}$$

According to usual terminology, we call the rows of $\mathbf{A}_{Row}$ as codewords of *outer row codes* or columns of $\mathbf{A}_{Col}$ as codewords of *outer column codes.* In the case of PC encoding all the outer codes have equal parameters. In GCC case parameters of outer codes must be different. After the second step all rows and columns belong to the *inner code* (row or column) with the parameters n - length, and k - dimension.

3 Basic relations

By (1), it holds that

$$\mathfrak{I} = \mathbf{\Phi}\mathbf{C}\mathbf{\Phi}^{\cdot}$$

Let $\mathbf{V}$ be a word for decoding and matrix $\mathbf{E}$ be an error-matrix such that

$$\mathbf{V} = \mathbf{C} + \mathbf{E}.$$

Clearly,

$$\mathbf{\Phi}\mathbf{V}\mathbf{\Phi} = \mathbf{\Phi}(\mathbf{C}+\mathbf{E})\mathbf{\Phi} = \mathfrak{I} + \mathbf{\Phi}\mathbf{E}\mathbf{\Phi}$$

Two-dimensional transform $\mathbf{\Phi}\mathbf{E}\mathbf{\Phi}$ of error-matrix gives us all the syndromes for outer codes (on the positions independent of matrix $\mathfrak{I}$), simultaneously. Very important: these syndromes deal with the projections of a row or column error of $\mathbf{E}$.

Row or Column Fourier Transform gives the following result

$$\mathcal{E}_{Col} = \mathbf{\Phi}\mathbf{E} = \begin{bmatrix} \mathcal{E}_{k\times n} \\ \mathcal{E}_{n-k\times n} \end{bmatrix}, \quad \mathcal{E}_{Row} = \mathbf{E}\mathbf{\Phi} = \left[\mathcal{E}_{n\times k}|\mathcal{E}_{n\times n-k}\right],$$

where $\mathcal{E}_{n-k\times n}$ and $\mathcal{E}_{n\times n-k}$ are the syndroms of all column inner codes and all row inner codes, respectively, independent of information matrix $\mathfrak{I}$. Very important: syndromes of row and column inner codes deal with row and column errors in the error-matrix.

Now we can conclude, that after the first (Row or Column) Fourier Transform we can decode all inner row or column codes alternately and correct some of errors in $\mathbf{E}$; after the second Fourier Transform we are able to decode all the outer codes and find a part of information matrix $\mathfrak{I}$, but more important to decode only the last carrent layer of GCC and make the transition to subcodes.

4 Decoding algorithms

For simplisity we use here new notations: $\mathcal{C}$ for any column codes and $\mathcal{R}$ for any row codes.

Standard GCC decoding

1. Set the starting layer of GCC $i = \ell$.

2. Inner code $\mathcal{C}_i$ decoding in all columns (results: correction or rejection).

3. Outer code $\mathcal{R}_i$ decoding in the i-th row (result: correction or rejection).

4. If the result of Step 3 is rejection then Stop decoding, else transition to the $i - 1$ -th layer, $i = i - 1 > 0$, of GCC and go to Step 2.

Iterative PC decoding (simple algorithm)

1. Column code $\mathcal{C}$ decoding in all columns (results: correction and/or rejection).

2. Row code $\mathcal{R}$ decoding in all rows (results: correction and/or rejection).

3. If no one correction and/or rejection were made on the Step 2 then stop decoding, else return to Step 1 until *additional stop condition* is not satisfied,

Iterative GCC decoding (general idea)

1. Set the starting layer of GCC $i = \ell$, $j = \ell$.

2. Iterative Product Code $\mathcal{R}_i \times \mathcal{C}_j$ decoding (results: correction and/or rejection).

3. Outer codes $\mathcal{R}_i$ and $\mathcal{C}_j$ decoding in i-th row and j-th column (result: correction and/or rejection).

4. If the result of Step 3 is rejection in the both row $\mathcal{R}_i$ and column $\mathcal{C}_j$ then Stop decoding, else transition to the $i = i - 1($ correction in the row) and /or $j = j - 1$ (correction in the column) layer of GCC and go to Step 2.

Introduce notations: $d_{col,i}$, $d_{row,i}$ are code distances of inner column and row component codes, $D_{col,i}$, $D_{row,i}$ are code distances of outer component codes of GCC.

Lemma 1. *Minimal Stop-Set of GCC i-th layer is any configuration with* $d_{col,i}/2$ *errors in* $D_{row,i}$ *columns.*

Lemma 2. *Minimal Stop-Set of PC is any configuration with* $d_{col}/2$ *errors in* d_{row} *columns such that there is no one row of less* $d_{row}/2$ *errors.*

Lemma 3. *Minimal Stop-Set of GCC i-th layer with using of symmetry Fourier designed GCC is as follows: any configuration with $d_{col,i}/2$ errors in $D_{row,i}$ columns and with $d_{row}/2$ errors in $D_{col,i}$ rows.*

It is evident that the stop-set of Lemma 3 is a subset of the stop-set of Lemma 1. A stop-set of Lemma 2 may be a part of stop-sets of Lemmas 1 and 3. If yes, then stop condition will satisfied for GCC decoding algorithms as for PC decoding. Anyway stop-set of the Lemma 3 is a non minimal stop-set for PC decoder.

Important property of PC stop-set in comparing with a stop-set of GCC is shown in the following lemma.

Lemma 4. *Not all non minimal stop-sets of GCC (standard or iterative) decoder include a PC decoder stop-set.*

Proof. Let us expand a minimal stop-set of Lemma reflem3 (for example) adding one row and one column. Then we get a configuration of $D_{row,i} + 1$ rows of the weight $d_{col,i}/2$ and $D_{col,i} + 1$ columns of the weight $d_{row,i}/2$. Remove any one error from this configuration. Then we get one row and one column of the weight 1 less than others and the rest of the configuration will be a minimal stop-set for GCC decoder. PC decoder iteratively and alternately finds and corrects the light column, then corrects all the light rows and so on up to correction all errors or to a stage when GCC decoder will be able to continue and finish correction of a rest of error configuration. □

There exists a number of different error configurations, correctable by iterative PC decoder with logarithmic of linear (of a configuration size) number of iterations.

The main result.

The considered configuration will stop both GCC decoding algorithms but, as a collorary of Lemma 4, the PC decoding can execute few successive iterations, after wards the GCC decoding will be able to finish the error correction with success. That means that union of PC and GCC decoding will expand a set of correctable error configurations.

Open problems:

- Estimation of improvements in probability of correct decoding or cardinality of correctable error configurations.
- Estimation of probability of decoding failure (error) because (it is clear) iterative procedure of PC decoding may increase an error propagation.

References

[1] R. E. Blahut, *Theory and Practice of Error Control Codes*, Addison-Wesley, Reading, 1984.

[2] E. L. Blokh, V. V. Zyablov, *Generalized Concatenated Codes*, Svyaz', Moscow 1976 (in Russian).

[3] E. L. Blokh, V. V. Zyablov, *Linear Concatenated Code*s, Nauka, Moscow 1982 (in Russian).

[4] I. Dumer, Concatenated codes and their multilevel generalizations, *Handbook of Coding Theory*, V. S. Pless and W. C. Huffman, Eds., Elsevier, Amsterdam, 1998, Ch. 23, 1911–1988.

[5] G. D. Forney , Jr., *Concatenated Codes*, MIT Press, 1966.

[6] V. A. Zinoviev, Generalized concatenated codes, *Problems Information Transmission*, **12**, 5–15, 1976.