Coded Interleaving for Burst Error Correction

L. R. Welch
University of Southern California

The concept of code interleaving has proved to be a very useful technique for dealing with complicated communications channels. One of the most recent applications of this concept is the Golay-Viterbi concatenation scheme proposed for use on the Mariner Jupiter/Saturn 1977 Mission. In this paper a generalization of interleaving is introduced. When two or more codes are suitably combined using this idea, the decoding algorithm for the first code can supply information about the location of errors for the remaining codes, thereby reducing the redundancy requirements for these codes.

I. Introduction

The concept of code interleaving has proved to be a very useful technique for dealing with complicated communications channels. One of the most recent applications of this concept is the Golay-Viterbi concatenation scheme proposed for use on Mariner Jupiter/Saturn 1977 (Ref. 1). In this paper a generalization of interleaving is introduced. When two or more codes are suitably combined using this idea, the decoding algorithm for the first code can supply information about the location of errors for the remaining codes, thereby reducing the redundancy requirements for these codes.

We begin with a somewhat abstract definition of interleaved codes. Let C_i be a block code of length N_i over an alphabet V_i, a vector space of dimension k_i over $GF(p)$.

Let

$$V = \sum_{i=1}^{M} \oplus V_i$$

the direct sum space whose elements are represented by M-tuples $(v_1, \cdots, v_m)$ where $v_i \in V_i$. V has dimension $k = \sum k_i$ over $GF(p)$.

If $\{c_i(t)\}$ is the sequence of symbols formed by encoding a sequence of symbols from source S_i using code C_i, then the *interleaved code* is the block code of length $N = LCM(N_1, \cdots, N_M)$ whose code words are of the form

$$\{c(t)\}_1^N = \{(c_1(t), \cdots, c_m(t))\}_1^N$$

The code is *interleaved to depth M.*

When the size of the symbol alphabet of a code C is different from the size of the channel alphabet, coding of the alphabet is necessary. When the channel has p symbols and the code alphabet is a k-dimensional vector space over $GF(p)$, the coding can be achieved by selecting a basis for the vector space and transmitting the k coefficients $(a_1, \cdots, a_k)$ of a vector relative to the basis. The resulting sequence of kN terms from $GF(p)$ will be called the channel code representing C. For example, if $V_1, \cdots, V_m$ are 1-dimensional in the above definition of interleaved codes and the basis for V is

$$\{(1,0, \cdots, 0), (0,1,0, \cdots, 0), \cdots, (0, \cdots, 0, 1)\}$$

then the channel code representing C is the usual definition of interleaved codes. That is, the subsequence $\{a(j), a(j+k), a(j+2k) \cdots\}$ is just a sequence obtained from the jth code used in the interleaving. In general, let $U_1, \cdots, U_k$ be a basis for V and

$$c(t) = \sum_{i=1}^{k} a_i(t)\, U_i$$

where $(c(1), \cdots, c(N))$ is a codeword in C. The set of kN-tuples $\{(a_1(1), \cdots, a_k(1), a_1(2), \cdots, a_k(N))\}$ is the *coded interleaved* code formed from $\{C_1, \cdots, C_M\}$.

II. Analysis

The question arises as to the choice of basis for V. This depends on the channel error statistics and the mode of decoding.

Let P_i be the projection of V onto V_i; that is

$$P_i(v_1, \cdots, v_m) = v_i$$

and let $u_1, \cdots, u_k$ be a basis for V. Then the mapping

$$(a_1, \cdots, a_k) \xrightarrow{L_i} \sum_j a_j P_i u_j$$

is a linear mapping from the set of k-tuples to V_i and the M-tuple $(L_1, \cdots, L_m)$ is the inverse of the coded interleaving mapping. If the component codes are to be decoded independently from a received sequence $\{r_i(t)\}$, then clearly the sequence $v_j(t) = L_j(r_1(t), \cdots, r_k(t))$ should be the input to the decoder for the jth code. The criterion for the choice of basis should be to minimize the error rate in as many component codes as possible. This choice depends on channel statistics. For example, let $(a_1(t), \cdots, a_k(t))$ and $(\bar{a}_1(t), \cdots, \bar{a}_k(t))$ be the only possibilities for $(r_1(t), \cdots, r_k(t))$, where bar denotes complementation. Then the basis should be chosen so that $L_j(1, \cdots, 1) = 0$ for all but one choice of j. M-1 codes would then be error free.

Another channel extreme occurs when

$$\text{Prob}\,((r_1 \cdots r_k) = (a_1, \cdots, a_k)) = 1 - P$$

$$\text{Prob}\,((r_1 \cdots r_k) = (b_1, \cdots, b_k)) = P/(p^k - 1)$$

$$\text{for } (b_1, \cdots, b_k) \neq (a_1, \cdots, a_k)$$

In this case, if an error occurs, it causes errors in each component code with high probability. Thus, if the decoding algorithm for one code detects an error in a given symbol, the corresponding forms in the other codes might just as well be erased and their decoding algorithm attempt to correct erasures.

The above example suggests that the following strategy may be useful for a large class of channels: use code C_1 as an error correcting code. If it is decided that symbol $c_1(t)$ is in error, erase $c_j(t)$ for $j = 2, \cdots, M$. Decode $C_2, \cdots, C_m$ as erasure channels.

It may happen that errors occur which do not affect C_1 and, therefore, are not erased in $C_2, \cdots, C_M$. The following theorem is useful in this context.

Theorem. *Let the probability of error in a channel have the property*

$$\text{Prob}\,[(r_1 \cdots r_k) = (a_1 \cdots a_k) + (e_1 \cdots e_k)] = P(e_1 \cdots e_k)$$

For each $k_1 < k$ there exists a linear transformation L_1 from k-tuples onto $V^{k_1}(p)$ for which

$$P_{L_1} = \text{Prob}\,[L_1(e_1, \cdots e_k) = \mathbf{0} \mid (e_1 \cdots e_k) \neq (0, \cdots, 0)] \leq \frac{1}{p^{k_1}}$$

Proof: Let $\delta(\mathbf{0}) = 1$ and $\delta(v) = 0$ for $0 \neq v \in V^{k_1}(p)$

$$P = \text{Prob}\,[L_1(e_1, \cdots, e_k) = \mathbf{0} \mid (e_1, \cdots, e_k) \neq (0, \cdots, 0)]$$

$$= \frac{1}{1 - P(0, \cdots, 0)} \cdot \sum_{(e_1 \ldots e_k) \neq 0} P(e_1 \cdots e_k) \times \delta(L_1(e_1, \cdots, e_k))$$

If we choose a basis for V^{k_1} and for the space of k-tuples, L_1 is represented by a k_1 by k matrix whose k_1 rows are linearly independent. There are

$$(p^k - 1)(p^k - p) \cdots (p^k - p^{k_1 - 1})$$

such matrices. We average the above equation over all such linear mappings.

$$\prod_{j=0}^{k_1-1} (p^k - p^j)^{-1} \sum_{L_1} P_{L_1} =$$

$$\frac{1}{1 - P(0, \cdots, 0)} \cdot \sum_{(e_1 \ldots e_k) \neq 0} P(e_1, \cdots, e_k)$$

$$\times \sum_{L_1} \prod_{j=0}^{k_1-1} (p^k - p^j)^{-1} \delta (L_1 (e_1, \cdots, e_k))$$

The innermost sum is nonzero only if $L_1 (e_1, \cdots, e_k) = 0$. If we choose as our basis for k-tuples one such that $(e_1, \cdots, e_k)$ is the first basis vector then L_1 with $L_1 (e) = 0$ is described by a k_1 by k matrix whose first column is zero and whose rows are linearly independent. The number of such is

$$\prod_{j=0}^{k_1-1} (p^{k-1} - p^j)$$

The above equation then reduces to

$$\text{Ave}(P_{L_1}) = \frac{1 - P(0, \cdots, 0)}{1 - P(0, \cdots, 0)} \cdot \frac{\prod_{j=0}^{k_1-1} (p^{k-1} - p^j)}{\prod_{j=0}^{k_1-1} (p^k - p^j)}$$

$$= \frac{1}{p^{k_1}} \cdot \frac{p^k - p^{k_1}}{p^k - 1} \leq \frac{1}{p^{k_1}}$$

Since the average of P_{L_1} is less than or equal to $1/p^{k_1}$, there exists an L_1 such that $P_{L_1} \leq 1/p^{k_1}$.

Q.E.D.

The theorem can be applied to the above decoding strategy as follows: If P_E is the probability of a k-tuple error in the channel and C_1 is capable of correctly decoding all errors "seen" by it. Then the other component codes will have an erasure rate of P_E and an additional error rate of $P_E \cdot 1/p^{k_1}$. For reasonable values of k_1 the other codes need be capable of finding few errors in addition to the erasures.

A variant of the above is that the extra errors found by C_2 be used to insert extra erasures in C_3, etc. The design problem is not of finding the best L_1 then the best L_2, etc., but finding the best sequence $(L_1, L_2, \cdots, L_M)$.

III. Conclusion

The concept of coded interleaving provides a richer class of codes than simple interleaving. It provides the ability to match the interleaving process to the channel statistics, thus allowing lower redundancy codes.

Reference

1. Baumert, L. D., and McEliece, R. J., "A Golay-Viterbi Concatenation Scheme," in this issue.