NONBINARY CODES CORRECTING LOCALIZED ERRORS

R. Ahlswede, L.A. Bassalygo, M.S. Pinsker

Abstract

A recursive construction yields for all alphabets asymptotically dense codes correcting a constant number of localized errors.

1. Introduction and main result

The authors of the paper [1], where codes correcting localized errors were introduced, noticed that an extension of results to nonbinary alphabets causes difficulties. In particular this is the case with the asymptotical Hamming bound, when the number of errors $t = \tau n$ increases linearly with the length n of the code. Up to now we proved it only in a small fixed interval $[0, \tau_o]$, although we believe that Hamming's bound is the true bound in the intervall $[0, \frac{1}{2}]$ (if $\tau \geq \frac{1}{2}$, then the rate of transmission is equal to 0 because it is impossible to transmit even two messages, when $t \geq \frac{n}{2}$). However our joint efforts bore fruit for other, rather accurate, asymptotical estimates on the code size, when the number of errors t is constant. The binary case was studied in [2], but the method used there doesn't give the tight answer for nonbinary cases. Here we present another method.

During the transmission of q–ary words of length n over the channel at most t errors occur, and the encoder knows the set E of t positions, where these errors are possible. The decoder doesn't know anything about these positions. Let $\mathcal{E}_t = \{E \mid E \subseteq \{1, 2, \ldots, n\}, |E| = t\}$ be the set of all subsets from $\{1, 2, \ldots, n\}$ of size t and let $\mathcal{M}$ be a set of messages $(|\mathcal{M}| = M)$. A code word $x(m, E)$ depends not only on the message $m \in \mathcal{M}$ but also on the configuration of possible errors E . So there exists the natural correspondence between the message $m \in \mathcal{M}$ and the list of code words $\bigcup_{E \in \mathcal{E}_t} \{x(m, E)\}$, which we use for the transmission of this message. Thus the code X for the set of messages $\mathcal{M}$ represents a collection of M lists $\left\{\bigcup_{E \in \mathcal{E}_t} \{x(m, E)\}, m \in \mathcal{M}\right\}$. Since we can use the same word for different configurations, the size of a list can be essentially smaller than the size of the set $\mathcal{E}_t (|\mathcal{E}_t| = \binom{n}{t})$.

Let us define the cylinder $\mathcal{C}(\mathfrak{a}, A)$ with the base $\mathfrak{a} = (a_1, \ldots, a_n)$ and the support A ($A \subseteq \{1, 2, \ldots, n\}$) as the set of words $(y_1, \ldots, y_n)$ with $y_i = a_i$, if $i \notin A$. It is clear that the size of the cylinder $\mathcal{C}(\mathfrak{a}, A)$ is equal to $q^{|A|}$ and the number of different cylinders with the same support A is equal to $q^{n-|A|}$.

As a result of the transmission of the codeword $x(m, E)$ every word of $\mathcal{C}(x(m, E), E)$ can appear as output of the channel. The code X corrects t localized errors, if the decoder can correctly recover every message $m \in \mathcal{M}$. The following condition is necessary and sufficient for it:

$$\mathcal{C}\big(x(m, E), E\big) \cap \mathcal{C}\big(x(m', E'), E'\big) = \varnothing \text{for all } \; E, E' \in \mathcal{E}_t, \; m, m' \in \mathcal{M}, m \neq m'.$$

The maximal number of messages, which we can transmit by a code correcting t localized errors, is denoted by $L_q(n, t)$.

Proposition 1.

$$L_q(n, t) \leq \frac{q^n}{S_t}$$

where $S_t = \sum_{i=0}^{t} (q-1)^i \binom{n}{i}$ is the size of a sphere of radius t in the Hamming n–space.

A proof of this bound in the q–ary case can be given as for the binary case in [1] or [3]. The key inequality there has the following generalization.

Lemma 1. *Let $\mathcal{C}(\mathfrak{a}_1, A_1), \ldots, \mathcal{C}(\mathfrak{a}_T, A_T)$ be cylinders with pairwise different supports $A_i \neq A_j, i \neq j$. Then for the size of the union of the cylinders*

$$\Big|\bigcup_{i=1}^{T} \mathcal{C}(\mathfrak{a}_i, A_i)\Big| \geq \sum_{i=1}^{T} (q-1)^{|A_i|}.$$

The proof of this lemma can be made by an induction on n as in [1] or [3] (for other proofs see [4]).

It yields immediately

$$\Big|\bigcup_{E \in \mathcal{E}_t} \mathcal{C}\big(x(m, E), E\big)\Big| \geq \binom{n}{t} (q-1)^t.$$

However, it can be used more efficiently. Choose any monotone map $f : \bigcup_{i=1}^{t} \mathcal{E}_i \to \mathcal{E}_t$ with the property $F \subset f(F)$ for all $F \in \bigcup_{i=1}^{t} \mathcal{E}_i$. Then we can write

$$\mathcal{C}\big(x(m, E), E\big) = \bigcup_{F \subset E} \mathcal{C}\big(x(m, f(F)), F\big).$$

Thus by Lemma 1