6.895 Essential Coding Theory	October 20, 2004
Lecture 11	
Lecturer: Madhu Sudan	*Scribe: Anastasios Sidiropoulos*

1 Overview

This lecture is focused in comparisons of the following properties/parameters of a code:

- List decoding, vs distance.
- Distance, vs rate.
- List decoding, vs rate.

2 The Plotkin's Bound

Recall that for two binary strings $x, y \in \{0,1\}^n$, we denote by $\Delta(x, y)$ the number of positions that x and y differ.

Theorem 1 (Plotkin's Bound) *If there exist codewords $c_1, c_2, \ldots, c_m \in \{0,1\}^n$, such that for each i, j, with $i \leq j$, $\Delta(c_i, c_j) \geq n/2$, then $m \leq 2n$.*

Proof Assume that $m > 2n$. We define vectors $\tilde{c}_1, \ldots, \tilde{c}_m \in \{-1, 1\}^n \subset \mathbb{R}^n$, such that for each i, with $1 \leq i \leq n$, $\tilde{c}_i$, and for each i, with $1 \leq j \leq n$, the jth coordinate of $\tilde{c}_i$ is -1, iff the jth bit of c_i is 1. Note that if $\Delta(c_i, c_j) \geq n/2$, then this implies $\langle \tilde{c}_i, \tilde{c}_j \rangle \leq 0$. Intuitively, this means that if two codewords c_i, and c_j have large Hamming distance, then the angle between the corresponding vectors $\tilde{c}_i$, and $\tilde{c}_j$, should be large.

Pick a random unit vector $x \in \mathbb{R}^n$. We have that w.h.p., $\langle x, \tilde{c}_i \rangle \neq 0$, for all i, with $1 \leq i \leq m$. Moreover, since there are m codewords, either x, or $-x$ has strictly positive inner product with at least $m/2$ of the $\tilde{c}_i$s. We can assume w.l.o.g., that this holds for x. Since $m > 2n$, it follows that there exist $n+1$ vectors having strictly positive inner product with x. W.l.o.g., assume that these are the vectors $\tilde{c}_1, \ldots, \tilde{c}_{n+1}$.

Observe that a set of $n+1$ vectors in an n-dimensional space, cannot be linear independent. Thus, we can assume that there exist $\lambda_1, \ldots, \lambda_{n+1} \in \mathbb{R}$, with $\lambda_i > 0$, for each i, with $1 \leq i \leq j$, and $\lambda_i \leq 0$, for each i, with $j < i \leq n+1$, such that

$$\sum_{i=1}^{j} \lambda_i \tilde{c}_i - \sum_{i=j+1}^{n+1} \lambda_i \tilde{c}_i = 0$$

Define the vector $z = \sum_{i=1}^{j} \lambda_i \tilde{c}_i$. We have to consider the following two cases for z:

Case 1, $z \neq 0$: We have $\langle z, z \rangle > 0$. On the other hand,

$$\begin{aligned} \langle z, z \rangle &= \left\langle \sum_{i=1}^{j} \lambda_i \tilde{c}_i, \sum_{i=j+1}^{n+1} \lambda_i \tilde{c}_i \right\rangle \\ &= \sum_{i \leq j, i' > j} \lambda_i \lambda_{i'} \langle c_i, c_{i'} \rangle \\ &\leq 0 \end{aligned}$$

Thus, we obtain a contradiction.

Case 2, $z = 0$: We have

$$\sum_{i=1}^{j} \lambda_i \tilde{c}_i = 0,$$

and thus

$$\begin{aligned}
\langle z, x \rangle &= \left\langle \sum_{i=1}^{j} \lambda_i \tilde{c}_i, x \right\rangle \\
&= \sum_{i=1}^{j} \lambda_i \langle \tilde{c}_i, x \rangle \\
&> 0
\end{aligned}$$

The last inequality follows from the fact that $\lambda_i > 0$, for $1 \leq i \leq j$, and that $\langle \tilde{c}_i, x \rangle > 0$. This however is a contradiction, since $z = 0$, which implies that $\langle z, x \rangle = 0$.

■

3 The Johnson's Bound

Theorem 2 (Johnson's Bound) *For any ϵ, with $0 < \epsilon < 1$, if C is a $\left[n, ?, \left(\frac{q-1}{q}\right)(1-\epsilon)n\right]_q$-code, then C corrects less than $\left(\frac{q-1}{q}\right)(1-\sqrt{\epsilon})n$ errors, with lists of size $(q-1)n$.*

We will give a proof of Theorem 2, for the special case of $q = 2$.

Proof We will prove the contrapositive. That is, we assume that there exist $r, c_1, \ldots, c_m \in \{0,1\}^n$, such that for each i, with $1 \leq i \leq m$,

$$\Delta(r, c_i) \leq \frac{1-\tau}{2} n,$$

and for each $i \neq j$,

$$\Delta(c_i, c_j) \geq \frac{1-\epsilon}{2} n.$$

Define vectors $\tilde{r}, \tilde{c}_1, \ldots, \tilde{c}_m \in \{0,1\}^n \subset \mathbb{R}^n$, as in the proof of Theorem 1. We have that for each i, with $1 \leq i \leq m$,

$$\langle \tilde{r}, \tilde{c}_i \rangle \leq \tau n,$$

and for each $i \neq j$,

$$\langle \tilde{c}_i, \tilde{c}_j \rangle \geq \epsilon n.$$

We want to show that is $\tau > \sqrt{e}$, then $m \leq n$.

We have that the projection of each $\tilde{c}_i$ into r is "large", and that the angle between each pair of $\tilde{c}_i$, $\tilde{c}_j$ is also "large". Intuitively, the main idea of the proof is that these two properties cannot be satisfied simultaneously, if the number of the vectors $\tilde{c}_i$ is too large. We will verify this argument by considering the vectors $\tilde{c}_i - \alpha r$, for carefully chosen α, and show that the angle between each pair of such vectors is at least $90°$. Thus, we will obtain a bound on the number of such vectors.

Formally, we have

$$\begin{aligned} \langle c_i - \alpha r, c_j - \alpha \rangle &= \langle c_i, c_j \rangle - \alpha \langle c_i, r \rangle - \alpha \langle c_j, r \rangle + \alpha^2 \langle r, r \rangle \\ &\leq (\epsilon - 2\alpha\tau + \alpha^2) n \end{aligned}$$

By setting $\alpha = \sqrt{\epsilon}$, we obtain that the inner product between each pair of vectors $\tilde{c}_i - \alpha r$, and $\tilde{c}_j - \alpha r$ is

$$2\sqrt{\epsilon}(\sqrt{\epsilon} - \tau) n$$

Thus, for any $\tau < \sqrt{\epsilon}$, the inner product is negative, and the assertion follows by applying the Plotkin's Bound. ■

We note that for the case $q > 2$, the proof of Theorem 2 becomes more technical. More specifically, one needs to map each bit of a codeword c_i, into more than one coordinates of the corresponding vector $\tilde{c}_i$. For example, if we have codewords in $\{0, 1, 2\}^n$, we can map each symbol of a vector in $\mathbb{R}$, such that the angle between each vector is at lest $90°$.

4 Relating R with δ

4.1 Improving the Singleton Bound

Lemma 3 *If there exists a $(n, k, d)_2$-code, then there also exists a $(2d, k + 2d - n, d)_2$-code.*

Proof Let C be a $(n, k, d)_2$-code. C contains 2^k codewords, of length n. Thus, if we project each codeword into the first $n - 2d$ coordinates, there are at least 2^{k+2d-n} codewords, that are mapped into the same string. Since all these 2^{k+2d-n} codewords have the same prefix of length $n - 2d$, and since their distance is at least d, it follows that their pairwise distance in the last $2d$ bits should be at least d. Thus, the suffixes of these codewords form a $(2d, k + 2d - n, d)_2$ code. ■

It follows by Lemma 3 that for any $(n, k, d)_2$-code, with $k + 2d - n \leq \log 4d$, we have

$$R + 2\delta - 1 \leq 0.$$

4.2 The Elias-Bassalygo Bound

The main argument in the proof of the Hamming bound is that if we have k non-intersecting balls of radius $\frac{d-1}{/}2$, in $\{0, 1\}^n$, then the sum of their volumes cannot exceed 2^n. We will show how to extend this idea in the case of intersecting balls, by bounding the overlap.

Assume that we have a binary code of distance $\frac{1-\epsilon}{2}$. For each codeword $c \in \{0, 1\}^n$, we consider the ball in $\{0, 1\}^n$ of radius $\frac{1-\sqrt{\epsilon}}{2}$ around c. We have

$$2^k \text{Vol}\left(n, \frac{1 - \sqrt{\epsilon}}{2} n\right) \leq n 2^n,$$

and thus

$$2^{Rn} 2^{H\left(\frac{1-\sqrt{\epsilon}}{2}\right) n} \leq 2^{n + o(n)}.$$

This implies

$$R + H\left(\frac{1 - \sqrt{\epsilon}}{2}\right) \leq 1$$

So, if $\delta = \frac{1-\epsilon}{2}$, then $R + H\left(\frac{1}{2} - \frac{1}{2}\sqrt{1 - 2\delta}\right) \leq 1$.

4.3 The Case $\delta \to 0$

An interesting question is what are the best possible codes, when $\delta \to 0$. The Hamming bound gives

$$\begin{aligned} R &\leq 1 - H\left(\frac{\delta}{2}\right) \\ &\approx 1 - \frac{1}{2}(1+o(1))\delta \log_2 \frac{1}{\delta}. \end{aligned}$$

On the other hand, we know that there exist codes satisfying

$$R \geq 1 - (1+o(1))\delta \log_2 \frac{1}{\delta}.$$

4.4 The Case $\delta \to 1/2$

Another interesting question is what is the best possible value for R, in the case where $\delta = \frac{1-\epsilon}{2}$, with $\epsilon \to 0$. The Plotkin bound gives $R \leq 2\epsilon$. Also, the EB-bound gives $R = O(\epsilon)$.

On the positive side, we can show (even for the case of random codes), that there exist codes with $R = \Omega(\epsilon^2)$.

We also note that the Linear-Programming bound gives $R = \tilde{O}(\epsilon^2)$ (also known as MRRW-bound, or JPL-bound).

5 Relating R with p

We would like to know what is the best possible values for R, and p, such that for infinitely many n, we have $(n, Rn, ?)_2$-codes, that are (pn, n)-error-correcting.

The Shannon bound gives

$$R \leq 1 - H_2(p)$$

We will next prove that this bound is tight.

Lemma 4 *There exist codes, satisfying $R \geq 1 - H_2(p)$.*

Before we state the proof, we note that the same result can be obtained by using random codes in $\{0,1\}^n$, but the proof is rather technical.

Proof We will show that there exists a linear code of rate R, that is $(pn, n+1)$-error-correcting. We begin with an empty basis for the code, and we repeatedly increase the basis, by greedily adding one base-vector at a time.

More specifically, assume that we have already added the vectors $b_1, b_2, \ldots, b_k \in \{0,1\}^n$ in the basis. Let $C_i = \text{span}(b_1, \ldots, b_i\}$. We pick b_{i+1}, so as to minimize the value Φ_{i+1}, where for each i, the value Φ_i is given by the following potential function:

$$\Phi_i = \mathbf{E}\left[2^{|B(x,pn)\cap C_i}\right],$$

where the expectation is taken over the random choices of x, when x is distributed uniformly in $\{0,1\}^n$.

We have

$$\mathbf{E}[\Phi_{i+1}] \leq \Phi_i^2$$

Thus, we can conclude that there exist base vectors $b_1, \ldots, b_k$, such that

$$\Phi_k \leq \Phi_0^{2^k}$$

Note that

$$\Phi_0 = 1 + \frac{\mathrm{Vol}(n, pn)}{2^n}$$

To be continued in the next lecture . . . ■