Telecom Paris

Telecom202b, Coding Theory
September 2021

Assignment 1

Exercise 1 (RAID, distributed storage)**.** Redundant Arrays of Independent Disks consist of a set of disks such that any subset of s disks can be disabled and the others are still able to reconstruct any requested file (the system can tell which disks are disabled).The rate of a RAID system corresponds to the rate at which data is stored.

1. Design a RAID system for 7 disks and $s = 2$. To do this you may want to consider the $(7, 4)$ Hamming code.

2. What happens if we use this code and try correct 3 erasures?

Exercise 2. Let C be a code with minimum distance d. Prove that C can correct any pattern of e_1 errors and e_2 erasures provided that $2e_1 + e_2 + 1 \leq d$. (Hint: given an erasure pattern, consider the code obtained by the deleting the erasure positions.)

Exercise 3 (Perfect codes)**.** A code is a perfect $t-$error correcting code if the set of $t-$spheres centered on the codewords fill the Hamming space $\{0, 1\}^n$ without overlapping. Here we will show that such codes do not, in general, achieve the capacity of the BSC.

Consider a set of three codewords of length n. Let un denote the number of positions where the first codeword differs from both the second and the third codewords, let vn denote the number of positions where the second codeword differs from both the first and the third codewords, let wn denote the number of positions where the third codeword differs from both the first and the second codewords, and finally let zn denote the number of positions where the three codewords agree.

1. Argue that we can assume, without loss of generality, that one of them is the all-zero codeword.

2. Assuming that the code is $f \cdot n$-error correcting, give necessary conditions on u, v, w.

3. Show that for a certain range of f we must have $u + v + w > 1$ which is impossible.

4. Conclude that, for a certain range of f, perfect codes do not exist.

5. Reconcile this result with the Shannon's result which says that 'with high probability it is possibe to correct $f \cdot n$ errors with exponentially many codewords'.

Exercise 4 ($A(n, d, w), A(n, d)$)**.** For any integers n, d, w let $A(n, d, w)$ be the largest possible size of a set of binary vectors of length n and weight w whose minimum distance is at least d, and let $A(n, d)$ be the largest possible size of a set of length n binary vectors whose minimum distance is at least d. Prove that

$$A(n, d) \leq \sum_{w=0}^{n} A(n, d, w)$$