IERG6120 Coding for Distributed Storage Systems	Lecture 21 - 01/12/2016

Fractional repetition codes

Lecturer: Kenneth Shum *Scribe: Cao Qi*

Definition 1. *For positive integer θ, we denote the set $\{1, 2, \ldots, \theta\}$ by $[\theta]$. A* hypergraph *is a pair $(\mathcal{P}, \mathscr{B})$, where $\mathcal{P} = [\theta]$ is a set of points, for some positive integer θ, and $\mathscr{B} \subseteq 2^{\mathcal{P}}$ is a collection of subsets of $\mathcal{P}$. A hypergraph is also called a* set system*, and a hyperedge in $\mathscr{B}$ is called a* block*.*

Definition 2. *A hypergraph $(\mathcal{P}, \mathscr{C})$ is a called an (n, α, ρ)* fractional repetition *(FR) code if*

- *There are n hyperedges in $\mathscr{C}$;*
- *Each hyperedge in $\mathscr{C}$ has size α;*
- *For all $p \in [\theta]$, p appears exactly ρ times among the hyperedges in $\mathscr{C}$.*

If the point set $\mathcal{P}$ is understood from the context, we will simply say that $\mathscr{C}$ is an (n, α, ρ) FR code.

The code parameter θ is uniquely determined by the following relation

$$\theta\rho = \alpha n. \tag{1}$$

FR codes can used as an inner code of a concatenated coding scheme, which distributes κ information symbols over a finite field to n storage nodes. The outer code is a (θ, κ) MDS code. The size of finite field is assumed to be large enough such that an (θ, κ) MDS code exists. The inner is an (n, α, ρ) FR code. Each node is assigned a hyperedge in the FR code. For $i = 1, 2, \ldots, n$, we let $\mathcal{E}_i$ be the hyperedge in $\mathscr{C}$ that is associated with node i,

$$\mathscr{C} = \{\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n\}.$$

The encoding consists of two steps.

- The κ information symbols are first mapped to a codeword $(c_1, c_2, \ldots, c_\theta)$ of the outer MDS code.
- For $i = 1, 2, \ldots, n$, node i stores the coded symbols c_j for $j \in \mathcal{E}_i$.

It is clear that, by the construction of the inner FR code, each node stores exactly α coded symbols, and each coded symbol is replicated ρ times among the n nodes.

Definition 3. *The* code rate *of the FR codes is defined as*

$$R_{\mathscr{C}}(k) \triangleq \min_{\substack{I \subseteq [n] \\ |I| = k}} \Big| \bigcup_{i \in I} \mathcal{E}_i \Big|.$$

If we pick an MDS code with dimension $\kappa = R_{\mathscr{C}}(k)$, then any k among the n storage nodes contain sufficiently many distinct coded symbols to recover the κ information symbols. If a node fails, the content of the failed node can be recovered by downloading the missing symbols from some other storage nodes. The helping nodes only need to read out the required code symbol and send it to the new node. This mode of repair is called "uncoded repair", or "repair by transfer". The number of connects made by the new node is no more than α.

Example 1 The Fano plane is shown in Fig 1.

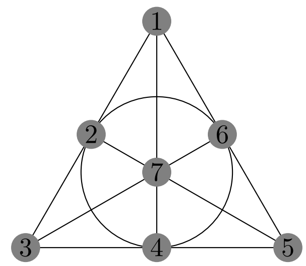


Figure 1: Fano plane

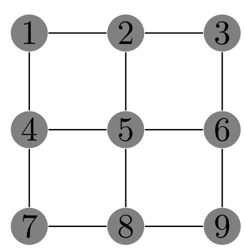
There are 7 points in this graph, and each point denotes one symbol. There are 7 lines in this graph, and each line denotes one node. Since every line consists of 3 points and every point is in three distinct lines, we have $\theta = 7$, $\mathcal{P} = \{1, 2, \ldots, 7\}$ and $\rho = 3$, and we can also obtain the FR code

$$\mathscr{C} = \{\{1,2,3\}, \{1,6,5\}, \{3,4,5\}, \{1,4,7\}, \{2,5,7\}, \{3,6,7\}, \{2,5,6\}\}.$$

Considering any two distinct lines have one point in common, now we try to construct the outer code according k.

- If $k = 2$, then any $k = 2$ lines contain $3 + 2$ distinct symbols. Thus $R_C(2) = 5$, i.e., the outer code is a $(7, 5)$ MDS code.
- If $k = 3$, then any $k = 3$ lines contain $3 + 2 + 1$ distinct symbols. Thus, the outer code is a $(7, 6)$ parity check code.

Example 2



There are 9 points in this graph, and each point denotes one symbol. There are 6 lines in this graph, and each line denotes one node. Since every line consists of 3 points and every point is in two distinct lines, we have $\theta = 9$, $\mathcal{P} = \{1, 2, \ldots, 9\}$, and $\rho = 2$. Thus, this is a $(6, 3, 2)$ FR code

$$C = \{\{1,2,3\}, \{4,5,6\}, \{7,8,9\}, \{1,4,7\}, \{2,5,8\}, \{3,6,9\}\}.$$

Now we try to construct the outer code. If $k = 3$, then any $k = 3$ lines at least contain $3 + 3 + 1 = 7$ distinct symbols, as is shown in the figures below.

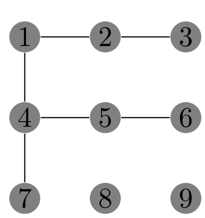


Figure 2: case 1

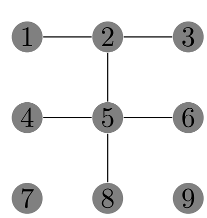


Figure 3: case 2

Thus, the outer code is a $(9,7)$ MDS code. Using the terminology of regenerating codes (lecture 15), the concatenated coding scheme is a distributed storage system with parameters $\beta = 1$, $d = 3$ and $k = 3$. According to the mincut bound, or what we call the min-sep bound in lecture 15, the maximal supported file size is upper bounded by $B \le 3 + 2 + 1 = 6$, which is strictly less than the file size 7 in this example. However, there is no contradiction. It is because the repair model in regenerating codes assume that a fail node can be repaired by any d other nodes. However, in this example, we cannot pick any d nodes to repair a failed node. For example, if $\{1,2,3\}$ fails, then we only can choose $\{1,4,7\}, \{2,5,8\}, \{3,6,9\}$ to repair. In FR-based distributed storage system, we can only guarantee that a failed node can be recovered by *some* choices of d other surviving nodes.

Theorem 4. *Let $\mathscr{C}$ be an (n, α, ρ) FR code, and k is given. We have*

$$R_{\mathscr{C}}(k) \le \left\lfloor \frac{n\alpha}{\rho} \left(1 - \frac{\binom{n-\rho}{k}}{\binom{n}{k}} \right) \right\rfloor.$$

Proof Let $\mathcal{E}_1, \ldots, \mathcal{E}_n$ be the hyperedges in $\mathscr{C}$. For a subset I of $[n]$, we define

$$U_I \triangleq \bigcup_{i \in I} \mathcal{E}_i.$$

Since $R_C(k)$ is the smallest $|U_I|$, with the minimum taken over index sets I of size k, we have

$$R_{\mathscr{C}}(k) \le \frac{1}{\binom{n}{k}} \sum_{\substack{I \subseteq [n] \\ |I| = k}} |U_I|.$$

The summation

$$\sum_{\substack{I \subseteq [n] \\ |I| = k}} |U_I|$$

on the right is equal to the number of pairs (p, I) that satisfies $p \in U_I$ and $|I| = k$. We can count this by first considering an arbitrary point $p_0 \in [\theta]$ and then count the number of index set I that satisfies $p_0 \in U_I$. Because the point p_0 belongs to ρ hyperedges in $\mathscr{C}$, the number of k-sets I such that $p_0 \notin U_I$ is equal to $\binom{n-\rho}{k}$. Thus,

$$\sum_{\substack{I \subseteq [n] \\ |I| = k}} |U_I| = \theta \left(\binom{n}{k} - \binom{n-\rho}{k} \right).$$

We obtain

$$R_{\mathscr{C}}(k) \le \left\lfloor \frac{\theta}{\binom{n}{k}} \left(\binom{n}{k} - \binom{n-\rho}{k} \right) \right\rfloor = \left\lfloor \frac{n\alpha}{\rho} \left(1 - \frac{\binom{n-\rho}{k}}{\binom{n}{k}} \right) \right\rfloor$$

as desired. □

In Example 2, we can use Theorem 4 to determine the bound of its code rate $R_{\mathscr{C}}(k)$. Since we have $k = 3$, $n = 6$, $\rho = 2$ and $\alpha = 3$, we would obtain

$$R_{\mathscr{C}}(k) \le \left\lfloor 9 \left(1 - \frac{\binom{4}{3}}{\binom{6}{3}} \right) \right\rfloor = 7.$$

Theorem 5. *Let $\mathscr{C}$ be an (n, α, ρ) FR code. Define $g(k)$ recursively by*

$$g(1) = \alpha$$

and

$$g(k+1) = g(k) + \alpha - \left\lceil \frac{\rho g(k) - \alpha k}{n-k} \right\rceil$$

for $k \geq 1$. Then we have

$$R_{\mathscr{C}}(k) \leq g(k)$$

for $k \leq n - \rho$.

Proof For $k = 1$, tt is obvious that $R_{\mathscr{C}}(1) = \alpha = g(1)$.

For the case $k = 2$, we fix an hyperedge, say $\mathcal{E}_1 \in \mathscr{C}$. We have the following upper bound

$$R_{\mathscr{C}}(2) \leq \frac{1}{n-1} \sum_{i=2}^{n} |\mathcal{E}_1 \cup \mathcal{E}_i|.$$

The cardinality of $\mathcal{E}_1 \cup \mathcal{E}_i$ is equal to 2α minus the number of points in the intersection $\mathcal{E}_1 \cap \mathcal{E}_i$. Hence

$$\sum_{i=2}^{n} |\mathcal{E}_1 \cup \mathcal{E}_i| = (n-1)2\alpha - \sum_{i=2}^{n} |\mathcal{E}_1 \cap \mathcal{E}_i|.$$

The summation on the right is equal to the number of pairs in the form (p, i), with $p \in [\theta]$ and $i \in \{2, 3, \dots, n\}$, such that $p \in \mathcal{E}_1 \cap \mathcal{E}_i$. For an arbitrary point $p_0 \in \mathcal{E}_1$, there are $\rho - 1$ indices i such that $p_0 \in \mathcal{E}_1 \cap \mathcal{E}_i$. Hence

$$\sum_{i=2}^{n} |\mathcal{E}_1 \cap \mathcal{E}_i| = \alpha(\rho - 1).$$

This gives

$$R_{\mathscr{C}}(k) \leq \left\lfloor \frac{1}{n-1} \Big(2\alpha(n-1) - \alpha(\rho-1) \Big) \right\rfloor = 2\alpha - \left\lceil \frac{\alpha(\rho-1)}{n-1} \right\rceil.$$

For the general case $k > 2$, we suppose that

$$R_{\mathscr{C}}(k) = |\mathcal{E}_1 \cup \mathcal{E}_2 \cup \cdots \cup \mathcal{E}_k|.$$

We can make this assumption without loss of generality by re-indexing the hyperedges in $\mathscr{C}$. We let $\mathcal{B}$ denote the union $\mathcal{E}_1 \cup \mathcal{E}_2 \cup \cdots \cup \mathcal{E}_k$. As in the previous cases, we can upper bound $R_{\mathscr{C}}(k+1)$ by

$$R_{\mathscr{C}}(k+1) \leq \frac{1}{n-k} \sum_{i=k+1}^{n} |\mathcal{B} \cup \mathcal{E}_i| = \frac{1}{n-k} \Big((|\mathcal{B}| + \alpha)(n-k) - \sum_{i=k+1}^{n} |\mathcal{B} \cap \mathcal{E}_i| \Big).$$

We evaluate the summation $\sum_{i=k+1}^{n} |\mathcal{B} \cap \mathcal{E}_i|$ by observing this is equal to the number of pairs $(p, i) \in [\theta] \times \{k+1, k+2, \dots, n\}$ satisfying $p \in \mathcal{B} \cap \mathcal{E}_i$. Since each point is replicated ρ times, there are totally $\rho|\mathcal{B}|$ pairs (p, j), with $j \in [n]$, such that $p \in \mathcal{E}_j$. Among these $\rho|\mathcal{B}|$ pairs, exactly αk of them does not satisfy $j \in \{k+1, k+2, \dots, n\}$. Therefore,

$$\sum_{i=k+1}^{n} |\mathcal{B} \cap \mathcal{E}_i| = \rho|\mathcal{B}| - \alpha k,$$

and thus

$$\begin{aligned} R_{\mathscr{C}}(k+1) &\leq |\mathcal{B}| + \alpha - \frac{\rho|\mathcal{B}| - \alpha k}{n-k} \\ &\leq g(k) + \alpha - \frac{\rho g(k) - \alpha k}{n-k}. \end{aligned}$$

The last line requires the assumption that

$$1 - \frac{\rho}{n-k} \geq 0,$$

which is equivalent to $n - k - \rho \geq 0$.

After taking the floor of both sides, we have

$$R_{\mathscr{C}}(k+1) \leq g(k) + \alpha - \left\lceil \frac{\rho g(k) - \alpha k}{n-k} \right\rceil.$$

The right-hand side is precisely the definition of $g(k+1)$. □

Definition 6. *A Steiner system is a set system* $([\theta], \mathscr{B})$*, denoted by* $S(t, \alpha, \theta)$*, satisfying*

- $|\mathcal{E}| = \alpha$ *for all* $\mathcal{E} \in \mathscr{B}$*;*
- *any* t *distinct symbols are contained in exactly one block in* $\mathscr{B}$*.*

Example 3 We consider $S(2, 3, 9)$, and the corresponding hypergraph is shown in the figure below.

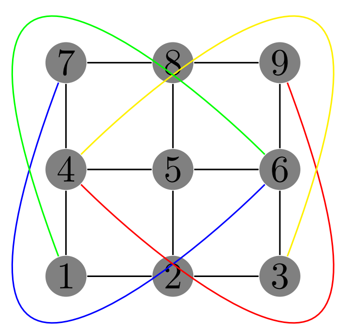


Figure 4: Affine plane

There are 9 points in this graph, and each point denotes one symbol. There are 10 lines in this graph, and each line denotes one node. We would find any $t = 2$ distinct symbols are contained in exactly one block. For example, $\{2, 5\}$ are only contained in {2,5,8}.

Since every line consists of $\alpha = 3$ points and every point is in four distinct lines, we have $\theta = 9$, $\mathcal{P} = \{1, 2, \ldots, 9\}$, and $\rho = 4$. Thus, we have a $(12, 3, 4)$ FR code

$$\begin{aligned}\mathscr{C} = \{&\{1,2,3\},\{4,5,6\},\{7,8,9\},\\ &\{1,4,7\},\{2,5,8\},\{3,6,9\},\\ &\{1,5,9\},\{3,4,8\},\{2,6,7\},\\ &\{3,5,7\},\{1,6,8\},\{2,4,9\}\}.\end{aligned}$$

Theorem 7. *In a Steiner system* $S(2, \alpha, \theta)$*, every point is repeated* ρ *times, where* $\rho = \frac{\theta-1}{\alpha-1}$*, and the total number of points is determined by* $\theta = \frac{\theta(\theta-1)}{\alpha(\alpha-1)}$*.*

Proof Let $\mathscr{B}$ be the collection of blocks in the Steiner system. For a point $p \in [\theta]$, let $r(p)$ denote the number of the blocks that contain p. We want to show that $r(p)$ is a constant regardless of the choice of p.

We fix a point $p_0 \in [\theta]$. We count the cardinality of the set

$$B(p_0) = \{(p, \mathcal{B}) \in [\theta] \times \mathscr{B} : \ p \neq p_0 \text{ and } \{p, p_0\} \subset \mathcal{B}\}.$$

Since each block contains α points, we would obtain $|B(p_0)| = (\alpha - 1)r(p_0)$. On the other hand, since each pair of points are contained in a unique block and there are $\theta - 1$ points in $[\theta] \setminus \{p_0\}$, we have $|B(p_0)| = \theta - 1$. Therefore, $r(p_0) = \frac{\theta-1}{\alpha-1}$, which is a constant independent of p_0.

Since $n\alpha = \theta\rho$ (1) the total number of blocks is equal to $n = \frac{\theta(\theta-1)}{\alpha(\alpha-1)}$. □

Corollary 8. *Given a Steiner system $S(2, \alpha, \theta)$, we can construct a $(\frac{\theta(\theta-1)}{\alpha(\alpha-1)}, \alpha, \frac{\theta-1}{\alpha-1})$ FR code $\mathscr{C}$ with rate*

$$R_{\mathscr{C}}(k) \geq k\alpha - \binom{k}{2}.$$

According to the code in Example 3, we would obtain the corresponding incidence matrix, as is shown below.

$$\underbrace{\left[\begin{array}{ccc|ccc|ccc|ccc}
1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
\hline
0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\
\hline
0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1
\end{array}\right]}_{blocks} \left.\vphantom{\begin{array}{c}1\\1\\1\\1\\1\\1\\1\\1\\1\end{array}}\right\} points$$

If we exchange the positions of blocks and points, we would obtain the transpose code, as is shown below.

$$\begin{aligned}
\mathscr{C}' = \{&\{1,4,7,11\}, \{1,5,9,12\}, \{1,6,8,10\}, \\
&\{2,4,8,12\}, \{2,5,7,10\}, \{2,6,9,11\}, \\
&\{3,4,9,10\}, \{3,5,8,11\}, \{3,6,7,12\}\}.
\end{aligned}$$

References

[1] S. El Rouayheb and K. Ramchandran, "Fractional repetition codes for repair in distributed storage systems," *Proc. 48th Annul Allerton Conf. Commun., Control, Comput. (Alleron)*, pp.1510–1517, Sep. 2010.