Distributed Storage Systems and Fractional Repetition Codes

Repair and Reconstruction Degree Analysis

Girish Patil, Manthan Patel
B. Tech. (ICT) – 2010

Abstract:

Given a fractional repetition code, compute and analyze the reconstruction and repair degree for the distributed storage system, using functional computer programs.

Introduction:

Distributed storage systems for large clusters typically use replication to provide reliability. Recently, erasure codes have been used to reduce the large storage overhead of three replicated systems. Reed-Solomon codes are the standard design choice and their high repair cost is often considered an unavoidable price to pay for high storage efficiency and high reliability. Distributed Storage Systems (DSSs) use coding theory to provide reliability in the system. We introduce Fractional Repetition code which consists of splitting the data of each node into several packets and storing multiple replicas of each on different nodes in the system.

Background:

i) Fractional Repetition Codes:

FR code is an arrangement of θ packets (each replicated ρ times in a smart way) on n nodes such that each node U_i, $1 \le i \le n$ has α_i packets.

Definition: A Fractional Repetition (FR) code denoted by $\varsigma(n, \vartheta, \alpha, \rho)$ *with replication factor* ρ*, for a DSS with parameter (n, k, d), is a collection* ς *of n subsets* $U_1, U_2, \dots, U_n$ *of a set* $\Omega = \{1, 2, \dots, \vartheta\}$*, which satisfies the following conditions:*

- *Every member of* Ω *appears exactly* ρ *times in the collection* ς*.*
- $|U_i| = \alpha_i$ $(i \in 1, 2, \dots, n)$
  where $\alpha = \max(\alpha_i)$, $i \in \{1, 2, \dots, n\}$

Clearly FR codes satisfy the equation

$$n\alpha = \rho\vartheta + \delta$$

where ϑ packets are replicated ρ times among n nodes and δ is total weakness of FR codes.

Table 1: Example of a node-packet distribution for FR code ς(7, 8, 4, 3):

Nodes	Packet Distribution	α_i	$\delta_i = \alpha - \alpha_i$	d_i
U_1	1, 6, 7, 8	4	0	2
U_2	1, 2, 7, 8	4	0	2
U_3	1, 2, 3 ,8	4	0	2
U_4	2, 3, 4, 7	4	0	2
U_5	3, 4, 5	3	1	2
U_6	4, 5, 6	3	1	2
U_7	5, 6	2	2	1

$\alpha = max\{ 4, 4, 4, 4, 3, 3, 2 \} = 4$, $\delta = \Sigma\delta_i = 4$ which satisfies the equation $n\alpha = \rho\vartheta + \delta$.

Node-packet distribution incidence matrix M_{7x8} for FR code $\varsigma(7, 8, 4, 3)$ is:

$$\begin{matrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{matrix}$$

ii) Equivalent Sequence

Definition: A sequence such that the packets at one or more nodes are arranged in a different manner, i.e. uniquely permuted, is called an equivalent sequence for given input sequence. Equivalent sequences retain the same amount and type of data, only the manner of representation differs.

Example:

Say, the input sequence is given by S = 123321, and $\vartheta = 2$.

S_1 = {1, 3}, S_2 = {2, 3}, and S_3 = {3, 1}. Now there are two possible permutations of S_1, {1, 3} and {3, 1}. Therefore, there are 2 possible equivalent sequences using the permutations of S_1, ①23③21 and ③23①21. Similarly, other equivalent sequences can be estimated by permuting the order of packets at one or more modes in the distributed system.

iii) Reconstruction Degree

Definition: The number k_{FR}, defines the minimum number of nodes to be contacted to retrieve $\vartheta - 1$ packets, to reconstruct the whole file from n nodes.

Example:

In the given FR code $\varsigma(7, 8, 4, 3)$, k^*= 2 as at least 7 packets can be recovered by contacting U_2 and U_5. Although, k_{FR} = 4, because we can recover the whole file i.e. at least $\vartheta - 1$ packets by contacting any 4 nodes in the distributed system.

iv) Repair Degree

Definition: The repair degree d_i, is defined for every node in the distributed system where $i \in \{1, 2, \ldots, n\}$, as the number of nodes to be contacted to recover the contents of that particular node in the case of a node failure.

Example:

The repair degrees of all nodes for the FR code $\varsigma(7, 8, 4, 3)$ are listed in Table 1. For instance, consider that node U_3 fails, and we need to revive the contents of the failed node. We can contact U_2 and U_1 to recover the packets at the failed node U_3.

Program 1:

i) Objective: To find all the equivalent sequences for a given input sequence.

ii) Inputs: Node-packet sequence, ϑ i.e. the number of packets.

iii) Outputs: ρ i.e. the replication factor and all the equivalent sequences.

iv) Algorithm:

a. Input node-packet sequence $S = x_1x_2x_3...x_n$, and ϑ.

b. Compute replication using the formula $\rho = n/\vartheta$.

c. Construct $S_1, S_2, S_3, \ldots, S_\vartheta$ such that $S_l = \{ l\vartheta + i \mid l \in Z, i \in \{1, 2, 3, .., \vartheta\} \}$

d. Permute S_i and increment $count_i$ *for each unique permutation.*

e. Display the packet sequence for all nodes, one-by-one.

f. Calculate the total number of unique permutation using the formula $\Pi\ count_i$ where $i \in \{1, 2, 3, .., \vartheta\}$.

Example:

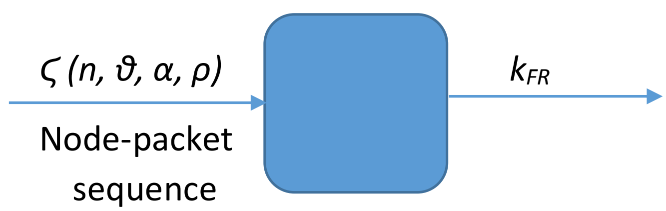
$S = 1234 \rightarrow n = 4$
$\vartheta = 2$
$n = \vartheta . \rho \rightarrow \rho = 2$
$S_1 = \{1, 3\}, S_2 = \{2, 4\}$
Permutations of $S_1 = \{1, 3\}, \{3, 1\} \rightarrow count_1 = 2$
Permutations of $S_2 = \{2, 4\}, \{4, 2\} \rightarrow count_2 = 2$

Product = $\Pi\ count_i$ = 2 x 2 = 4
Sequences = *1234, 1432, 3214, 3412*

Program 2:

i) Objective: To find the reconstruction degree for a given FR code.

Given a node-packet distribution for an FR code we need to find the minimum number of nodes required (k_{FR}) so that the entire data can be recovered. Suppose there are a total of θ packets including a parity packet which is added using the MDS code, one can always delete one packet from all the nodes (W.L.O.G., we usually delete last packet) as we can recover it using the parity of MDS codes. Hence for constructing entire data it is sufficient to reconstruct only (θ - 1) packets. Thus WLOG we delete last packet in the algorithm.



ii) Inputs: Node-packet distribution, n i.e. the number of nodes, ρ i.e. the replication factor, ϑ i.e. number of packets.

iii) Outputs: k_{FR}

iv) Algorithm:

Requires: A set of packets $\Omega = \{1, 2, 3,...., \vartheta\}$ and node-packet distribution of FR code with n nodes $U^n = \{U^1, U^2,.... , U^n\}$.

Ensure: Exact reconstruction degree $k_{FR.}$

1: For $1 < m < n$ set $U^m = \{U^1, U^2,.... , U^m\}$. Take $m = n.$

2: Pick the set $U_m \in U^m$ and call this set as P. Set the counter $k_\lambda = 1; 1 \leq k_m \leq m$ and $1 \leq \lambda \leq n$. If $\Omega/P = \phi$ or singleton set or singleton set then go to step 6 otherwise go to step 3.

3: If $\exists$ U_j $(1 \leq j \leq m) \in U^m$ s.t. $U_j \cap P = \phi$ then go to step 4 otherwise jump to step 5.

4: Pick an arbitrary $U_{j'}$ $(1 \leq j' \leq m) \in U^m$ which has maximum cardinality among all $U_{j'}$ in U^m with $U_{j'} \cap P = \phi$. Update $P = P \cup U_{j'}$, update counter $k_\lambda = (k_\lambda + 1)$. Again if $\Omega/P = \phi$ or singleton set then go to step 6 otherwise go to step 3.

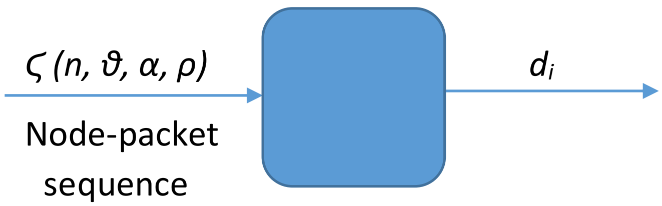
5: Pick U_r $(1 \leq r \leq m) \in U^m$ s.t. $U \not\subset P$ which has maximum $|Ur/P|$ among all $U_r \in U^m$ having the condition $U_r \not\subset P$ then update $P = P \cup U_r$, update counter $k_\lambda = (k_\lambda + 1)$. Again if $\Omega/P = \phi$ or singleton set then go to step 6 otherwise go to step 5.

6: Store k_λ in k'_λ and set $k_\lambda = k_{(\lambda+1)}$.

7: *If* $1 < \lambda < n$ then calculate $U^{m-1} = U^m \setminus \{U_m\}$ and perform step 2 for $P = U_{j''}$ $(1 \leq j'' \leq n) \in U^m$. Otherwise report $K_{FR} = max\ \{k'_\lambda\}^n{}_{\lambda = 1}$.

Program 3:

i) Objective: To find the repair degree for a given FR code.



ii) Inputs: Node-packet distribution, n i.e. the number of nodes, ρ i.e. the replication factor, ϑ i.e. number of packets.

iii) Outputs: d_i for all nodes.

iv) Algorithm:

a. Compute the incidence matrix $M_{nx\vartheta}$ of the given FR code.

b. For each node i, $1 \le i \le n$ let $S_i^{\{\,i\,\}} = \{\, H_j \setminus \{\, i \,\} \mid i \in H_j\,,\ 1 \le j \le \vartheta \,\}$

c. Compute $T \subseteq \{1, 2, \ldots, \vartheta\}$ s.t. $|T|$ is maximum among all possible subsets and for $t \in T$, $H_t \setminus \{\, i \,\} \in S_i^{\{\,i\,\}}$, and $\cap\ H_t \setminus \{\, i \,\} \neq \phi$. Set counter $l_q(1 \le q \le n) = |T| - 1.$

d. Update $S_i^{\{\,i\,\}} = S_i^{\{\,i\,\}} \setminus (\, H_t \setminus \{\, i \,\}\,),\ \ \forall t \in T.$

e. If $S_i^{\{\,i\,\}} = \phi$ then $d_i = \alpha_i - \Sigma l_\lambda$, where $\alpha_i = |V_i|$, otherwise set $q = q + 1$ and go to step 2.

Example:

For a given FR code ς *(n, ϑ, α, ρ)* and with parameters $n = 8$, $\vartheta = 8$, $\alpha = 4$, $\rho = 2$, the aforementioned algorithms produce the following output:

Nodes	**Packet Distribution**	K_{FR}	d_i
U_1	1, 2, 3, 4	3	2
U_2	1, 2, 5, 7	2	0
U_3	3, 4, 6, 8	2	3
U_4	7, 8	3	2
U_5	6	3	1

References: