

Adaptive-and-Resolvable Fractional Repetition Codes Based on Hypergraph

Tiantian Wang¹, Jing Wang^{2*}, Haipeng Wang¹, Jie Meng¹, Chunlei Yu³, and Shuxia Wang⁴

¹Naval Aviation University

Yantai, Shandong 264001 China

[e-mail: TTian_work@163.com, whp5691@163.com, 15564855530@163.com]

²School of Information Engineering, Chang'an University

Xi'an, Shaanxi 710064 China

[e-mail: jingwang@chd.edu.cn]

³Intelligent Manufacturing Institute, Sichuan University of Arts and Science

Dazhou, Sichuan 635000 China

[e-mail: yclinfinite@163.com]

⁴Shanxi Jinzhong Institute of Technology

Jinzhong, Shanxi 030600 China

[e-mail: 805326752@qq.com]

*Corresponding author: Jing Wang

*Received August 20, 2022; revised February 2, 2023; accepted March 15, 2023;
published April 30, 2023*

Abstract

Fractional repetition (FR) codes can achieve exact uncoded repair for multiple failed nodes, with lower computational complexity and bandwidth overhead, and effectively improve repair performance in distributed storage systems (DSS). The actual distributed storage system is dynamic, that is, the parameters such as node storage overhead and number of storage nodes will change randomly and dynamically. Considering that traditional FR codes cannot be flexibly applied to dynamic distributed storage systems, a new construction scheme of adaptive-and-resolvable FR codes based on hypergraph coloring is proposed in this paper. Specifically, the linear uniform regular hypergraph can be constructed based on the heuristic algorithm of hypergraph coloring proposed in this paper. Then edges and vertices in hypergraph correspond to nodes and coded packets of FR codes respectively, further, FR codes is constructed. According to hypergraph coloring, the FR codes can achieve rapid repair for multiple failed nodes. Further, FR codes based on hypergraph coloring can be generalized to heterogeneous distributed storage systems. Compared with Reed-Solomon (RS) codes, simple regenerating codes (SRC) and locally repairable codes (LRC), adaptive-and-resolvable FR codes have significant advantages over repair locality, repair bandwidth overhead, computational complexity and time overhead during repairing failed nodes.

Keywords: Fractional Repetition (FR) codes, distributed storage systems (DSS), adaptive-and-resolvable FR codes, hypergraph, heterogeneous distributed storage systems.

This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 62001059, the Key Research and Development Project of Shaanxi Province under Grant 2021GY-019.

1. Introduction

Nowadays, the information data is growing explosively. Due to significant characteristics of high throughput, high availability and high scalability, the large-scale distributed storage systems have been widely deployed to store massive data. However, node failures are common in large-scale distributed storage systems. Thus some redundancy strategies are usually adopted. The traditional redundancy strategies are replication mechanism [1] and erasure codes strategy [2]. The replication mechanism is to store multiple replicas of original file in different nodes, so storage overhead is too high. The erasure codes strategy requires high network bandwidth during repairing failed nodes, so repair bandwidth overhead is too high.

Dimakis et al. proposed the regenerating codes (RC), which can effectively reduce node storage overhead and repair bandwidth overhead [3]. Then Rashmi et al. proposed two classes of regenerating codes that achieve optimal limits—minimum storage regeneration (MSR) codes with optimal storage overhead and minimum bandwidth regeneration (MBR) codes with optimal repair bandwidth overhead [4]. Although reducing bandwidth overhead of repairing failed codes, the proposed regenerating codes don't take into account disk I/O overhead. The disk I/O overhead is directly proportional to the number of surviving nodes connected in repair process [5]. In order to ensure fewer nodes connected in repair process, Papailiopoulos and Dimakis proposed locally repairable codes (LRC) [6]. Combining RS codes with XOR operations, Papailiopoulos et al. proposed a class of simple locally repairable codes, termed as simple regenerating codes (SRC) [7].

However, regenerating codes and LRC involve a large number of operations over finite field $GF(q)$ in repair process, that is to say, the computational complexity in repair process is too high. To further reduce computational complexity of repairing failed nodes, El Rouayheb and Ramchandran proposed Fractional Repetition (FR) codes, as a new class of MBR codes [8]. The FR codes can tolerate low-complexity uncoded repair for multiple failed nodes, with minimum repair bandwidth overhead and disk I/O overhead, thus performance of repairing node failures is improved significantly.

The traditional FR codes are mainly applied to static distributed storage systems, which storage overhead and repetition degree of coded packets remain unchanged. However, due to some faults occur at any time in distributed storage systems, including partial data loss of storage nodes and storage node failures, the actual distributed storage system is dynamic [9]. Aiming at the problem that traditional FR codes cannot adapt to dynamic storage system changes, Zhu proposed adaptive FR codes and construction method based on symmetric designs [9]. When partial data of adaptive FR codes is lost, the remaining data still satisfies the properties of FR codes without reconfiguring the system, which saves a lot of bandwidth overhead [10]. Subsequently, Olmez and Ramamoorthy constructed FR codes based on combinatorial design, and proposed resolvable FR codes constructed based on resolvable designs [11-13]. When nodes fail in dynamic distributed storage systems, resolvable FR codes can changes repetition degree by simply adjusting surviving nodes. Considering the changes of storage overhead and repetition degree of coded packets in dynamic distributed storage systems, based on circulant permutation matrices (CPMs) and affine permutation matrices (APMs), Su proposed adaptive-and-resolvable FR codes [14]. For large-scale distributed storage systems, the construction scheme involves a large number of operations. The FR codes only adapt to dynamic homogeneous storage systems, which storage overhead of nodes and

repetition degree of coded packets are consistent. Thus the above FR codes cannot be applied to heterogeneous distributed storage systems.

For large-scale dynamic distributed storage systems, we propose a construction scheme of adaptive-and-resolvable FR codes based on hypergraph coloring in this paper. These FR codes constructed can be flexibly applied to dynamic distributed storage systems, where the node storage overhead, number of storage nodes, and repetition degree of coded packets can change randomly and dynamically. By proposing the existence condition of linear uniform regular hypergraph, this paper further confirms the existence condition of adaptive-and-resolvable FR codes. Moreover, the adaptive-and-resolvable FR codes can be easily constructed for any specified parameters in distributed storage system. Specifically, based on the heuristic algorithm of hypergraph coloring proposed in this paper, the linear uniform regular hypergraph is constructed. The edges and vertices in hypergraph correspond to nodes and coded packets of FR code respectively, further, the FR code is constructed. Due to different popularity of original data packets and coded data packets, based on hypergraph coloring, we generalize the construction of adaptive-and-resolvable FR codes to heterogeneous distributed storage systems. Theoretical analysis shows that, according to hypergraph coloring, the FR codes can achieve rapid repair for multiple failed nodes. Compared with RS codes, SRC and LRC, the FR codes have best performance in repair locality, repair bandwidth overhead, computational complexity and time overhead during repairing failed nodes.

In this work, we mainly introduce adaptive-and-resolvable FR codes based on hypergraph coloring. In Section II, we introduce the preliminaries, mainly including hypergraph and adaptive-and-resolvable FR codes. The construction scheme based on hypergraph coloring is proposed in section III. In section IV, we study rapid repair method based on hypergraph coloring for single failed node and multiple failed nodes. In Section V, we generalize the construction of FR codes based on hypergraph coloring to heterogeneous distributed storage systems. Theoretical analysis and simulation results are shown in section VI. In Section VII, we draw a conclusion.

2. Preliminaries

2.1 Hypergraph

Definition 1 (hypergraph) [15]: As hypergraph $H = (V, E)$, V is a finite set of vertices in hypergraph, and E is a set of non-empty subsets in V called edges in hypergraph. An edge can connect at least one vertex.

The hypergraph with no loops or multiple identical edges is termed as simple hypergraph [16]. The linear hypergraph is a simple hypergraph with at most one common vertex in any two edges. If each edge contains r vertices, the hypergraph is termed as r -uniform hypergraph. Similarly, if each vertex is contained by t edges, the hypergraph is termed as t -regular hypergraph. If linear hypergraph is both r -uniform hypergraph and t -regular hypergraph, the hypergraph is termed as linear r -uniform t -regular hypergraph, i.e. (r, t) -hypergraph [16].

Definition 2 (incidence matrix): A hypergraph H with vertices set $V = \{v_1, v_2, \dots, v_n\}$ and edges set $E = \{e_1, e_2, \dots, e_m\}$ has a $n \times m$ incidence matrix $A = (a_{ij})_{n \times m}$, where $a_{ij} = 1$ if $v_i \in e_j$, otherwise $a_{ij} = 0$.

For hypergraph H , if edges e_i and e_j ($i \neq j$) contain the same vertex, the edges e_i and e_j are adjacent edges. Similarly, if vertices v_i and v_j ($i \neq j$) are contained by the same edge, the vertices v_i and v_j are adjacent vertices.

Definition 3 (locally correlated edges): For hypergraph H , assumed that edges e_i and e_j ($i \neq j$) only contain vertices v_i and v_j respectively, if vertices v_i and v_j are adjacent vertices, edges e_i and e_j are termed as locally correlated edges.

2.2 Adaptive-and-Resolvable FR Codes

In (n, k, d) distributed storage systems, assume that the number of storage nodes is n and the node storage overhead is α , the original files can be reconstructed by connecting any k storage nodes. And d surviving nodes can be connected to repair a single failed node. The above parameters satisfy $k \leq d \leq n-1$ and $d = \alpha$.

Definition 4 (FR codes) [8]: In (n, k, d) distributed storage systems, FR code $C = (\Omega, N)$ is a set of n subsets, i.e. $N = \{N_1, \dots, N_n\}$, the symbols in each subset belong to symbol set $\Omega = \{1, \dots, \theta\}$, and repetition degree of C is ρ [17]. The (n, d, θ, ρ) FR code has two characteristics.

(a) The size of subset N_i is d .

(b) Each symbol in Ω belongs to ρ subsets in set N , and each pair of subsets contain one common symbol at most.

According to above definition, subset N_i represents the storage node of FR code C , the symbol in set Ω represents the coded packet by MDS encoding, and the elements in subset N_i represents the coded packets stored in storage node N_i . For FR code C , the amount of coded packets is θ , the repetition degree is ρ , all coded packets are stored in n storage nodes, and the storage overhead of nodes is d . Therefore, the parameters of FR code satisfy $\theta\rho = nd$.

Definition 5 (Adaptive FR Codes) [9]: In FR code $C = (\Omega, N)$, if there is $S \subset \Omega$ and the set consisting of non-empty subsets $N_0 \setminus S, N_1 \setminus S, \dots, N_{n-1} \setminus S$ is an FR code, then C is an adaptive FR code.

In practical distributed storage systems, partial data loss of storage nodes is common phenomenon, especially the systems cannot be recovered for permanent failure. For adaptive FR code C , due to the remaining coded packets of storage nodes still satisfy the properties of FR codes, a new FR code C' can be obtained by simply adjusting FR code C . And FR code C' can be applied to new distributed storage system without reconfiguring the system.

Definition 6 (resolvable FR code): In FR code $C = (\Omega, N)$, as a subset $P \subset N$, if $N_i \in P$ and $N_j \in P$ ($i \neq j$), there exists $N_i \cap N_j = \emptyset$ and $\bigcup_{\{j: N_j \in P\}} N_j = \Omega$, then P is termed as a parallel class. If N consists of several parallel classes, then FR code C is termed as a resolvable FR code [13].

According to Definition 6, any two nodes in a parallel class do not contain the same coded packets, and all nodes in a parallel class contain all coded packets. There exist at least two parallel classes in resolvable FR code. The resolvable FR codes can dynamically change the

repetition degree by adjusting parallel classes to adapt to dynamic distributed storage systems. Combining Definition 5 with Definition 6, the FR codes with both adaptive and resolvable properties are termed as adaptive-and-resolvable FR codes [14].

3. Adaptive-and-Resolvable FR Codes Based on Hypergraph Coloring

Consider that the (d, ρ) -hypergraph $H = (V, E)$ with vertices set $V = \{v_1, v_2, \dots, v_\theta\}$ and edges set $E = \{e_1, e_2, \dots, e_n\}$, any two edges contain one same vertex at most.

The existence condition of (d, ρ) -hypergraph H is as follows:

- (1) $n \equiv 0 \pmod{\rho}$;
- (2) $n/\rho = \theta/d$ or $nd = \theta\rho$;
- (3) $d^2 \leq \theta$;
- (4) $\rho^2 \leq n$.

The heuristic algorithm for hypergraph coloring is as follows:

Step 1: The vertices set $V = \{v_1, v_2, \dots, v_\theta\}$ is divided into d vertices subsets by sequence, each subset contains θ/d vertices, which are $V_1 = \{v_1, \dots, v_{\theta/d}\}$, \dots , $V_d = \{v_{(\theta-\theta/d+1)}, \dots, v_\theta\}$ respectively. The edges set $E = \{e_1, e_2, \dots, e_n\}$ is divided into ρ edges subsets by sequence, which are $E_1 = \{e_{1,1}, \dots, e_{1,\theta/d}\}$, \dots , $E_\rho = \{e_{\rho,1}, \dots, e_{\rho,\theta/d}\}$ respectively.

Step 2: For $t=1$, the indices of edges $1 \leq j \leq \theta/d$ and the indices of vertices $1 \leq i \leq \theta$, we assign vertices to the edges $e_{1,1}, \dots, e_{1,\theta/d}$. When $i = j \pmod{\theta/d}$, the vertex $\{v_i \mid i = j \pmod{\theta/d}, 1 \leq i \leq \theta\}$ is assigned to the edge $e_{1,j}$.

Step 3: For $2 \leq t \leq \rho$ and the indices of edges $1 \leq j \leq \theta/d$, according to the vertices assigned to edges $e_{1,1}, \dots, e_{1,\theta/d}, \dots, e_{t-1,1}, \dots, e_{t-1,\theta/d}$, we assign vertices to edges $e_{t,1}, \dots, e_{t,\theta/d}$. This situation satisfies the following:

- (1) The edges $e_{t,1}, \dots, e_{t,\theta/d}$ contain all vertices and any two edges are not adjacent or locally correlated;
- (2) There is only one vertex of each vertices subset V_1, \dots, V_d in edge $e_{t,j}$ ($1 \leq j \leq \theta/d$);
- (3) Any two vertices in vertices set V can be contained by one edge of $e_{1,1}, \dots, e_{1,\theta/d}, \dots, e_{t-1,1}, \dots, e_{t-1,\theta/d}, e_{t,1}, \dots, e_{t,\theta/d}$ at most;

Step 4: Hypergraph coloring. For $1 \leq t \leq \rho$, the edges subset $\{e_{t,1}, \dots, e_{t,\theta/d}\}$ are assigned the same color. There are ρ different colors in the hypergraph, and each vertex in set $V = \{v_1, v_2, \dots, v_\theta\}$ is contained by ρ edges of different colors.

The edges and vertices in hypergraph correspond to the storage nodes and coded packets of above FR codes respectively. The vertices assigned to edges in hypergraph correspond to the coded packets stored in nodes of FR codes. The edges with the same color in hypergraph correspond to a parallel class in FR codes. The (d, ρ) -hypergraph is constructed based on the heuristic algorithm of hypergraph coloring. Further, (n, d, θ, ρ) adaptive-and-resolvable FR codes can be constructed.

Example 1: For $(4, 3)$ -hypergraph $H = (V, E)$, vertices set $V = \{v_1, v_2, \dots, v_{16}\}$ and edges set $E = \{e_1, e_2, \dots, e_{12}\}$. First, the vertices set V is divided into 4 vertices subsets by sequence, i.e. $V_1 = \{v_1, v_2, v_3, v_4\}$, $V_2 = \{v_5, v_6, v_7, v_8\}$, $V_3 = \{v_9, v_{10}, v_{11}, v_{12}\}$, and $V_4 = \{v_{13}, v_{14}, v_{15}, v_{16}\}$; the edges set E is divided into 3 edges subsets, i.e. $E_1 = \{e_1, e_2, e_3, e_4\}$, $E_2 = \{e_5, e_6, e_7, e_8\}$, and $E_3 = \{e_9, e_{10}, e_{11}, e_{12}\}$.

As shown in Fig. 1, according to the heuristic algorithm of hypergraph coloring, let vertices are assigned to all edge subsets. The vertices contained by edges $\{e_1, e_2, e_3, e_4\}$ are $\{v_1, v_5, v_9, v_{13}\}$, $\{v_2, v_6, v_{10}, v_{14}\}$, $\{v_3, v_7, v_{11}, v_{15}\}$, $\{v_4, v_8, v_{12}, v_{16}\}$ respectively. The vertices contained by edges $\{e_5, e_6, e_7, e_8\}$ are $\{v_1, v_6, v_{11}, v_{16}\}$, $\{v_2, v_5, v_{12}, v_{15}\}$, $\{v_3, v_8, v_9, v_{14}\}$, $\{v_4, v_7, v_{10}, v_{13}\}$ respectively. The vertices contained by edge $\{e_9, e_{10}, e_{11}, e_{12}\}$ are $\{v_1, v_7, v_{12}, v_{14}\}$, $\{v_2, v_8, v_{11}, v_{13}\}$, $\{v_3, v_5, v_{10}, v_{16}\}$, $\{v_4, v_6, v_9, v_{15}\}$ respectively. Finally, the same edges subset is assigned the same color, so the chromatic number of all edges is 3.

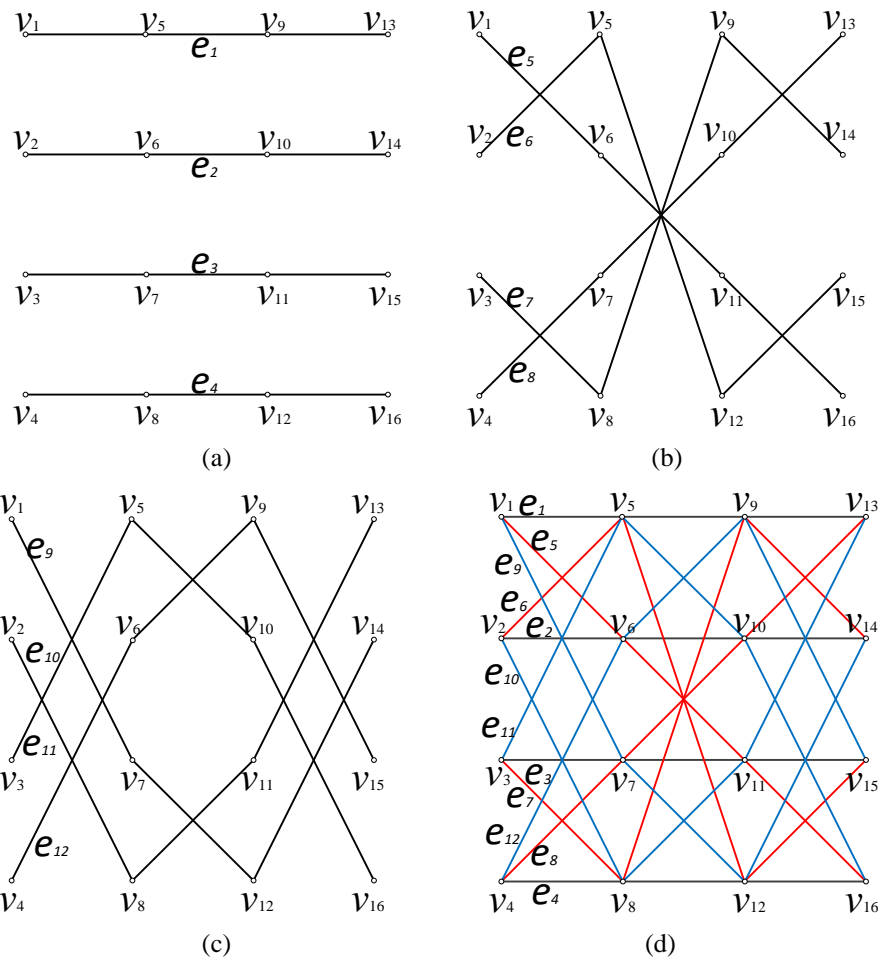


Fig. 1. Based on heuristic algorithm of hypergraph coloring, the (d, ρ) -hypergraph is constructed.

According to the (4, 3)-hypergraph constructed in Example 1, the incidence matrix A is acquired. The vertices of hypergraph correspond to row vectors of incidence matrix, and edges of hypergraph correspond to column vectors of incidence matrix.

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

The above incidence matrix can also represent FR code. The row vectors of incidence matrix correspond to coded packets of FR code, and column vectors correspond to storage nodes of FR code. Similarly, the weight of row vectors represents the repetition degree ρ of coded packets, and the weight of column vectors represents storage overhead d of nodes. Thus, the (12, 4, 16, 3) adaptive-and-resolvable FR code is constructed based on incidence matrix. As shown in Fig. 2, there are 3 parallel classes $\{N_1, N_2, N_3, N_4\}$, $\{N_5, N_6, N_7, N_8\}$ and $\{N_9, N_{10}, N_{11}, N_{12}\}$ in FR code, corresponding to 3 edges subsets $E_1 = \{e_1, e_2, e_3, e_4\}$, $E_2 = \{e_5, e_6, e_7, e_8\}$ and $E_3 = \{e_9, e_{10}, e_{11}, e_{12}\}$ in hypergraph.

| | | | | | | | | | | | | | | |
|-------|---|---|----|----|-------|---|---|----|----|----------|---|---|----|----|
| N_1 | 1 | 5 | 9 | 13 | N_5 | 1 | 6 | 11 | 16 | N_9 | 1 | 7 | 12 | 14 |
| N_2 | 2 | 6 | 10 | 14 | N_6 | 2 | 5 | 12 | 15 | N_{10} | 2 | 8 | 11 | 13 |
| N_3 | 3 | 7 | 11 | 15 | N_7 | 3 | 8 | 9 | 14 | N_{11} | 3 | 5 | 10 | 16 |
| N_4 | 4 | 8 | 12 | 16 | N_8 | 4 | 7 | 10 | 13 | N_{12} | 4 | 6 | 9 | 15 |

Fig. 2. The (12, 4, 16, 3) adaptive-and-resolvable FR code is constructed.

4. Repair of Failed Nodes Based on Hypergraph Coloring

According to hypergraph coloring situation, the adaptive-and-resolvable FR codes can repair the failed nodes.

4.1 Repair of Single Failed Node

For a single failed node, the newcomer node collects d coded packets from the surviving nodes corresponding to one complete edge subset (i.e. the surviving nodes in intact parallel class), which achieves the exact uncoded repair.

As shown in Fig. 3, when node N_1 corresponding to edge e_1 fails, the newcomer node can connect surviving nodes $\{N_5, N_6, N_7, N_8\}$ corresponding to complete edge subset $\{e_5, e_6, e_7, e_8\}$ to achieve the exact uncoded repair. Similarly, the newcomer node can also connect surviving nodes $\{N_9, N_{10}, N_{11}, N_{12}\}$ corresponding to complete edge subset $\{e_9, e_{10}, e_{11}, e_{12}\}$ to achieve the exact uncoded repair.

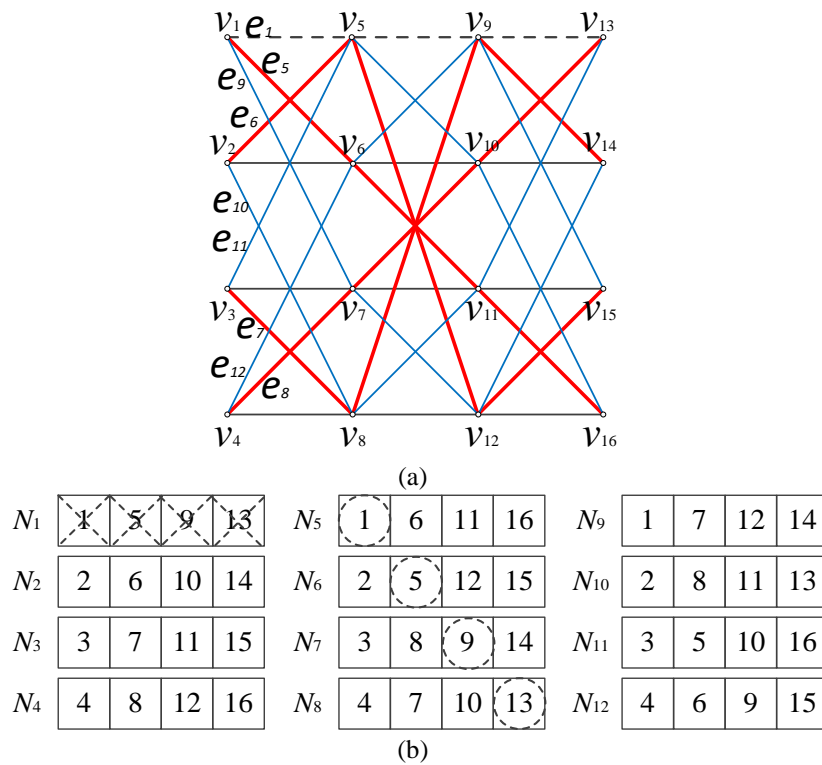


Fig. 3. Repair single failed node.

4.2 Repair of Multiple Failed Nodes

The ρ parallel classes in adaptive-and-resolvable FR code correspond to ρ chromatic edges subsets in hypergraph. The multiple failed nodes correspond to multiple edges in hypergraph. There are the following two cases of multiple failed nodes.

For the edges corresponding to multiple failed nodes with at most $\rho - 1$ colors (i.e. at least one complete edge subset exists in hypergraph), the newcomer nodes connect to surviving nodes corresponding to one complete edge subset to achieve the exact uncoded repair.

As shown in Fig. 4, when nodes N_1 and N_5 fail, the newcomer nodes can connect surviving nodes $\{N_9, N_{10}, N_{11}, N_{12}\}$ corresponding to complete edge subset $\{e_9, e_{10}, e_{11}, e_{12}\}$ to achieve the exact uncoded repair.

For the edges corresponding to multiple failed nodes with ρ colors and the number of failed nodes no more than $n - k$, if there are ρ failed nodes stored the same coded packet, the newcomer nodes can connect any k surviving nodes to reconstruct original file. If there are not ρ failed nodes stored the same coded packet, firstly find the vertices and edges corresponding to failed coded packets and failed nodes, then find the adjacent edges of the failed edges, finally the newcomer nodes can connect surviving nodes corresponding to adjacent edges to achieve exact uncoded repair.

As shown in Fig. 5, when nodes N_1, N_5 and N_{10} fail, the newcomer nodes can connect surviving nodes $\{N_2, N_3, N_7, N_8, N_9, N_{11}\}$ to achieve exact uncoded repair.

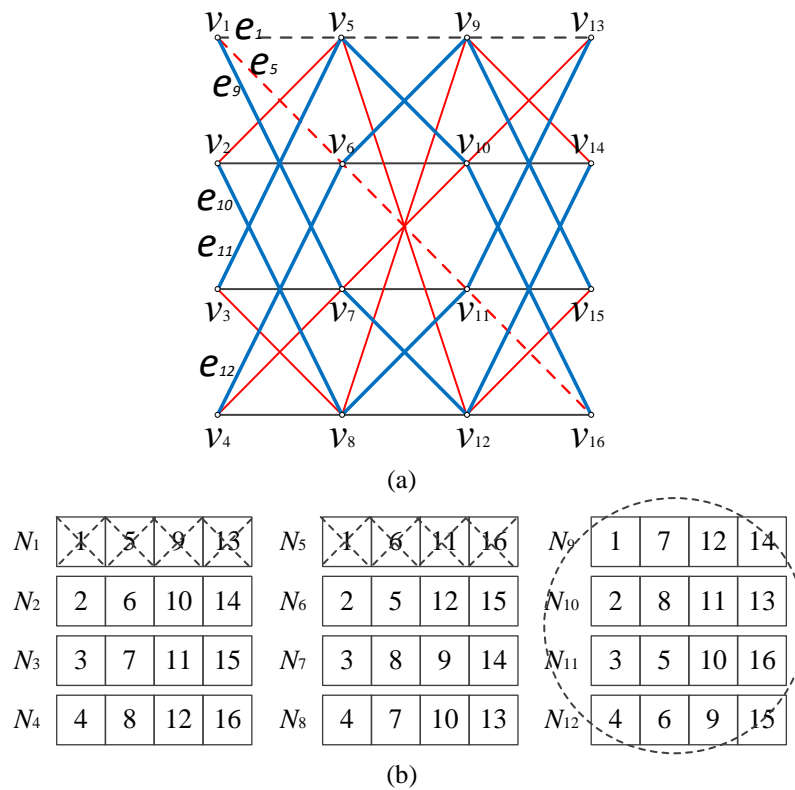


Fig. 4. For the edges corresponding to multiple failed nodes with at most $\rho - 1$ colors, repair multiple failed nodes.

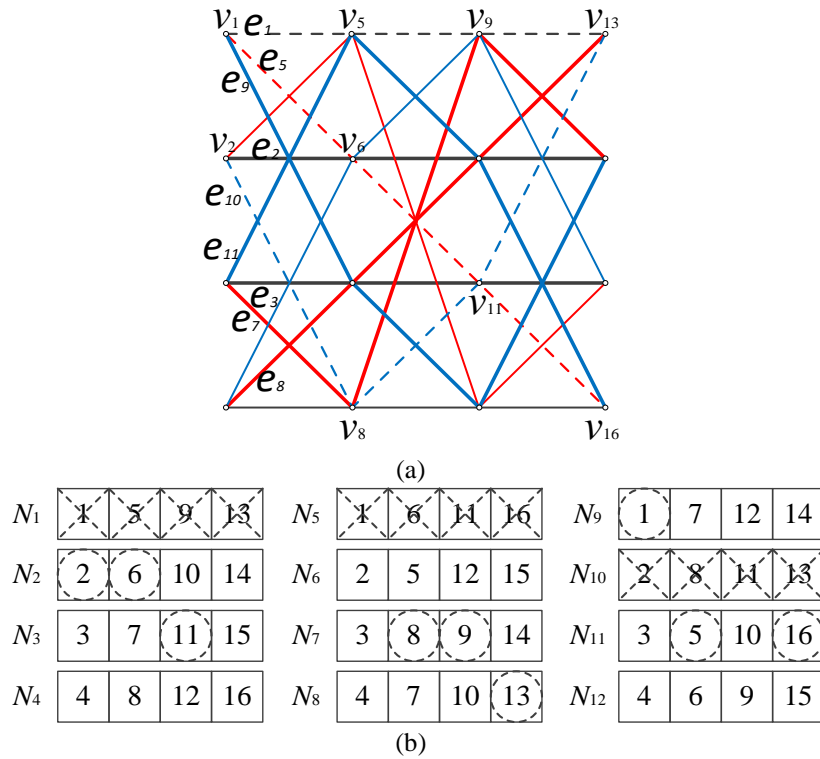


Fig. 5. For the edges corresponding to multiple failed nodes with ρ colors, the number of failed nodes no more than $n - k$, and no more than ρ failed nodes stored the same coded packet, repair multiple failed nodes.

5. Construction of FR Codes in Heterogeneous Storage Systems

We generalize the construction of FR codes from homogeneous storage systems to heterogeneous storage systems, where storage overhead d and repetition degree ρ can be different. Actually, the popularity of original data packets and coded data packets is different, it is desirable that the more popular packets will have higher repetition degrees in distributed storage systems. Thus, we propose a new idea based on hypergraph coloring, that the repetition degree of coded data packets is reduced when the repetition degree of original data packets remains unchanged. Based on (d, ρ) -hypergraph, by deleting the edges in hypergraph corresponding to coded packets of adaptive-and-resolvable FR code, the FR code can be simply extended to heterogeneous distributed storage system.

Consider that the (n, d, θ, ρ) FR codes with a set of storage nodes $N = \{N_1, \dots, N_n\}$ and a set of coded packets $\Omega = \{1, \dots, \theta\}$ with repetition degree ρ . Assume that connecting any k storage nodes to download m coded packets can reconstruct original file. Then the set of original data packets is $\Omega_1 = \{1, \dots, m\}$, the set of coded data packets is $\Omega_2 = \{m + 1, \dots, \theta\}$. According to the new idea proposed, the repetition degree ρ_1 of original data packets remains unchanged, that is $\rho_1 = \rho$, and the repetition degree ρ_2 of coded data packets is reduced, that is $\rho_2 < \rho$. Based on the heuristic algorithm of hypergraph coloring, when $\rho_2 < t \leq \rho$, we

delete the edges between the vertices $\{v_{m+1}, \dots, v_{\theta}\}$ and the other vertices, then FR codes in heterogeneous storage system can be constructed. For $(12, 4, 16, 3)$ adaptive-and-resolvable FR code with repetition degree $\rho = 3$ and $k = 4$, the set of original data packets is $\Omega_1 = \{1, \dots, 11\}$, and the set of coded data packets is $\Omega_2 = \{12, \dots, 16\}$. We specify that the repetition degree of original data packets is $\rho_1 = 3$ and that of coded data packets is $\rho_2 = 2$. Based on the heuristic algorithm of hypergraph coloring, the hypergraph is constructed as shown in Fig. 6(a), and the FR code in heterogeneous storage system is shown in Fig. 6(b).

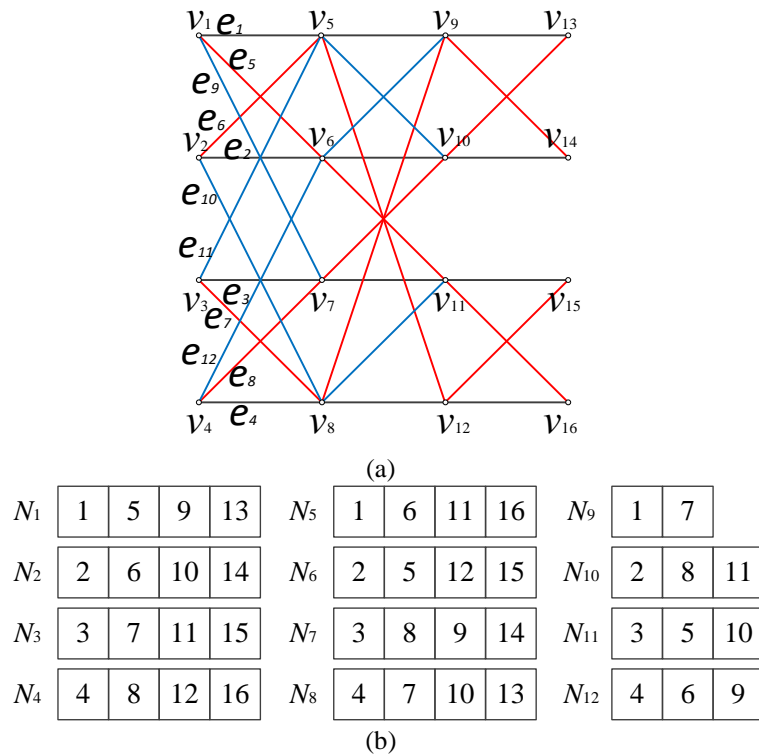


Fig. 6. The construction of FR codes is generalized from homogeneous storage systems to heterogeneous storage systems.

6. Performance Analysis

For adaptive-and-resolvable FR codes based on hypergraph, performance is mainly analyzed from three aspects: repair locality, bandwidth overhead, computational complexity and repairing time. Moreover, the performance is compared with the common RS codes, SRC and LRC. In distributed storage systems, it is assumed that the number of storage nodes is n , the size of original file is B , and the reconstruction degree to RS codes, SRC and LRC is k . In the section, we only consider the cases of single node and two nodes failures.

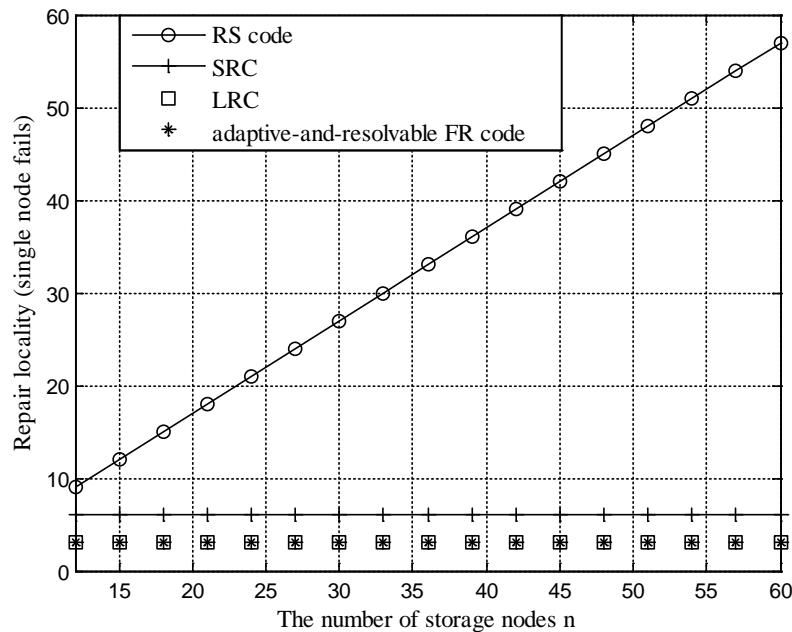
6.1 Repair Locality

When a single node fails, as (n, k) RS codes, k surviving nodes need to be connected, so the repair locality is k . As (n, k, f) SRC, the original file consists of f subfiles, each subfile is encoded by RS encoding, then $2f$ surviving nodes need to be connected to repair the failed

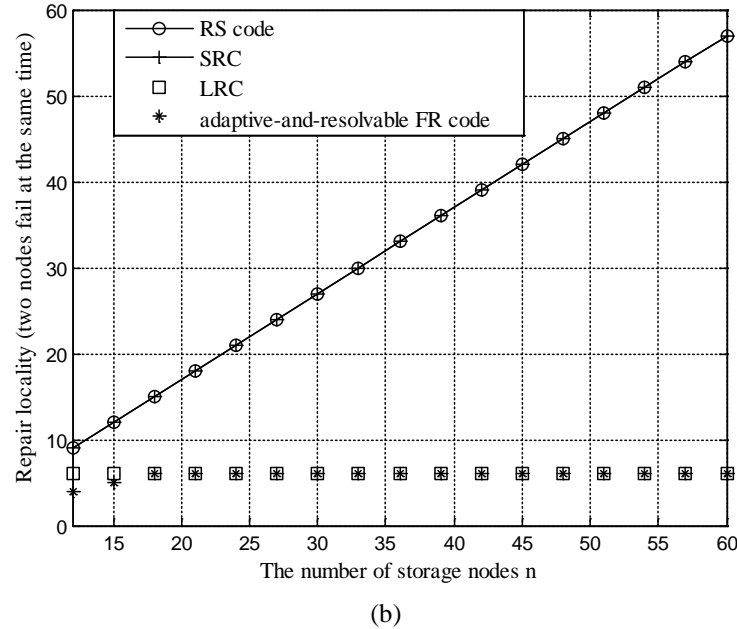
data, so the repair locality is $2f$. As LRC, the original file is divided into $r \times k$ data packets, then $r \times n$ coded packets are obtained by (n, k) RS encoding, and the n check packets are obtained by XOR operations of coded packets with the same subscript. The coded packets and check packets are stored in $r + 1$ nodes of each local repair group according to the subscript cycle. Then r surviving nodes in the local repair group need to be connected, so the repair locality is r . As adaptive-and-resolvable FR codes, d surviving nodes in a parallel class need to be connected, so the repair locality is d .

When two nodes fail simultaneously, as (n, k) RS codes and (n, k, f) SRC, the original file needs to be reconstructed by connecting k surviving nodes, so the repair localities are k . As LRC, when stored in different local repair groups, two failed nodes can be repaired, which is equivalent to a single node fails in each local repair group, so the repair locality is $2r$. As adaptive-and-resolvable FR codes with $\rho > 2$, there are two cases on repairing for two failed nodes. (1) For two failed nodes in same parallel class, obviously, two failed nodes don't store same coded packet, the $\min\{2d, n/\rho\}$ surviving nodes are connected of same parallel class. (2) For two failed nodes in different parallel classes, two failed nodes store at most one same coded packet, the $\min\{2d - 1, n/\rho\}$ surviving nodes are connected of same parallel class. When two nodes fail simultaneously, the maximum repair locality of adaptive-and-resolvable FR codes is $\min\{2d, n/\rho\}$.

Assuming that reconstruction degree of original file to (n, k) RS codes, (n, k, f) SRC and LRC is $k=n-3$. As (n, k, f) SRC, the original file consists of $f=3$ subfiles, and each subfile is divided into k data packets. Similarly, as LRC, the original file is divided into $3k$ data packets, that is to say, when a single node fails, the repair locality is $r=3$. For the convenience of performance analysis, adaptive-and-resolvable FR codes is externally encoded by (n, k) MDS, that is, the number of coded packets is $\theta=n$, and the repetition degree is $\rho=3$, so the node storage capacity is $d=3$. As shown in Fig. 7, whether a single node or two nodes fail, the adaptive-and-resolvable FR codes and LRC, have lower repair locality.



(a)



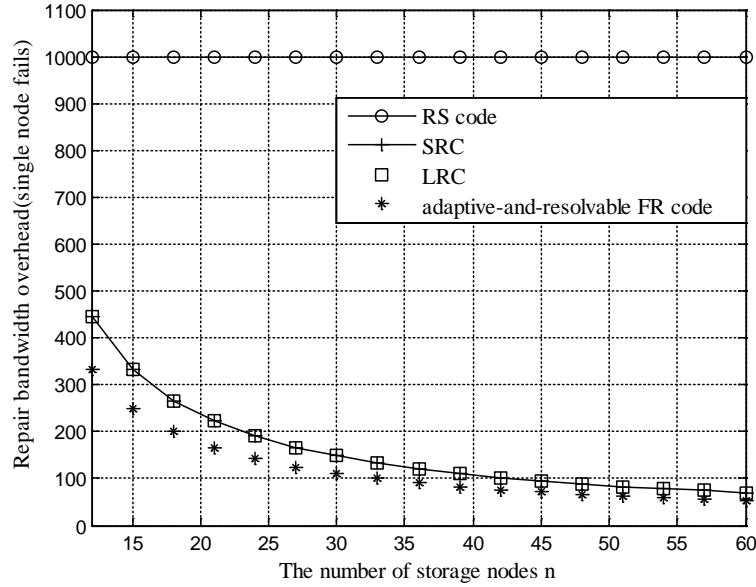
(b)
Fig. 7. The performance analysis is about repair locality.

6.2 Repair Bandwidth Overhead

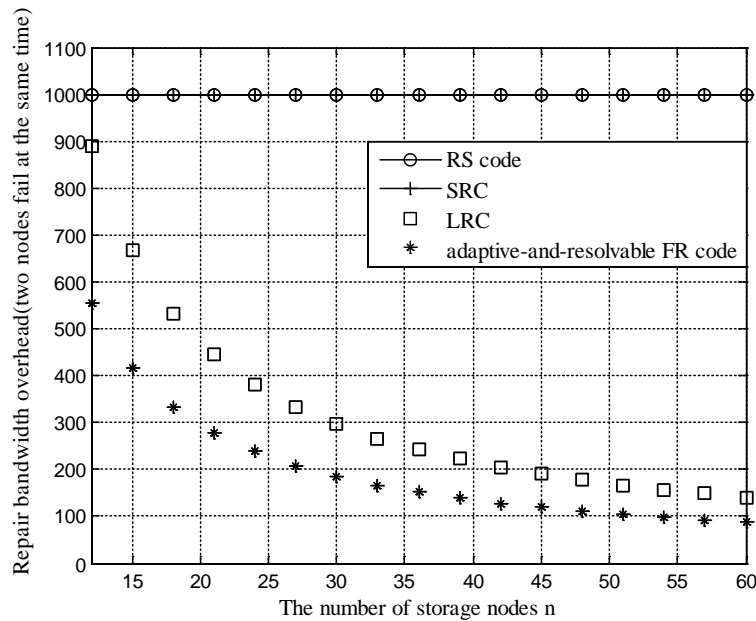
When a single node fails, as (n, k) RS codes, repair locality is k , and storage overhead of nodes is B/k , so repair bandwidth overhead is B . As (n, k, f) SRC, f coded packets need to be download to repair one coded packet, the storage overhead of nodes is $(f+1)B/fk$, so repair bandwidth overhead is $(f+1)B/k$. Similarly, as LRC, repair bandwidth overhead is $(r+1)B/k$. As adaptive-and-resolvable FR codes, the number of original data packets is j , the number of coded data packets is θ . When a single node fails, the d surviving nodes need to be connected, and one coded packet from each surviving node need to be downloaded, so repair bandwidth overhead is Bd/j .

When two nodes fail simultaneously, as (n, k) RS codes and (n, k, f) SRC, the original file needs to be reconstructed by connecting k surviving nodes, so their repair bandwidth overheads are B . As LRC, it is equivalent to repair one failed node in each local repair group, so repair bandwidth overhead is $2(r+1)B/k$. As adaptive-and-resolvable FR codes with $\rho > 2$, there are two cases on repairing for two failed nodes. (1) For two failed nodes in same parallel class, repair bandwidth overhead is $2Bd/j$. (2) For two failed nodes in different parallel classes, repair bandwidth overhead is $B(2d-1)/j$.

Assume that the size of original file is $B=1000\text{Mb}$, encoding parameters are described in above subsection. As shown in **Fig. 8**, whether a single node or two nodes fail, the adaptive-and-resolvable FR codes have lower repair bandwidth overhead.



(a)



(b)

Fig. 8. The performance analysis is about repair bandwidth overhead.

6.3 Computational Complexity and Repair Time

As (n, k) RS codes, during the process of repairing failed nodes, any k surviving nodes need to be connected to reconstruct original file, then coded packets of failed nodes can be recovered by encoding to original file. Each coded data packet is obtained by operations between k original data packets on finite field $GF(q)$. The decoding process of reconstructing original file is equivalent to encoding process. Thus the repair process of RS codes involves

$k^2 + k$ -times multiplication operations and $k^2 - 1$ -times addition operations over the finite field $\text{GF}(q)$. As (n, k, f) SRC, a coded packet can be repaired by $f - 1$ -times XOR operation, and each node stores $f + 1$ coded packets. Thus the repair process of SRC involves $(f - 1)(f + 1)$ -times XOR operations. Similar to SRC, the repair process of LRC involves $(r - 1)(r + 1)$ -times XOR operations. As adaptive-and-resolvable FR codes, the corresponding surviving nodes are connected based on repair table, the newcomer nodes can download the failed coded packets directly. Thus the repair process of the FR codes does not involve any operations over finite field $\text{GF}(q)$.

As shown in **Table 1**, adaptive-and-resolvable FR codes have the lowest computational complexity during repairing failed nodes.

Table 1. The computational complexity

| Encoded Methods | Computational Complexity |
|-------------------------------------|---|
| RS codes | $k^2 + k$ -times multiplication operations; $k^2 - 1$ -times addition operations |
| SRC | $(f - 1)(f + 1)$ -times XOR operations |
| LRC | $(r - 1)(r + 1)$ -times XOR operations |
| Adaptive-and-Resolvable FR codes | no operation |

When the failed nodes are repaired in distributed storage systems, RS codes involve a large number of operations over the finite field $\text{GF}(q)$, with high computational complexity and long repair time. SRC and LRC involve some simple XOR operations, which also increases computational complexity and long repair time to some extent. The uncoded repair process of adaptive-and-resolvable FR codes only involves reading for coded packets, so computational complexity is the lowest, and it greatly reduces the repair time of failed nodes.

7. Conclusion

The new construction scheme of adaptive-and-resolvable FR codes based on hypergraph coloring has been proposed in this paper. These FR codes constructed can be flexibly applied to dynamic distributed storage systems, where the node storage overhead, number of storage nodes, and repetition degree of coded packets change randomly and dynamically. In this paper, the existence condition of (d, ρ) -hypergraph and the heuristic algorithm of hypergraph coloring are proposed. According to specified parameters in distributed storage system, (d, ρ) -hypergraph and corresponding (n, d, θ, ρ) adaptive-and-resolvable FR can be constructed based on the heuristic algorithm of hypergraph coloring. Based on hypergraph coloring, this paper analyzes the rapid repair methods for failed nodes. Further, the FR codes can be generalized to heterogeneous distributed storage systems based on hypergraph coloring. Compared with RS codes, SRC and LRC, adaptive-and-resolvable FR codes have better repair locality and repair bandwidth overhead. Moreover, adaptive-and-resolvable FR codes have great advantages on computational complexity and time overhead during repairing failed nodes.

References

- [1] Y. Liu and V. Vlassov, "Replication in Distributed Storage Systems: State of the Art, Possible Directions, and Open Issues," in *Proc. of 2013 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery*, Beijing, China, pp. 225-232, 2013. [Article \(CrossRef Link\)](#)
- [2] J. Li and B. Li, "Erasure coding for cloud storage systems: A survey," *Tsinghua Science and Technology*, vol. 18, no. 3, pp. 259-272, Jun. 2013. [Article \(CrossRef Link\)](#)
- [3] A. G. Dimakis, P. B. Godfrey, Y. Wu, M. J. Wainwright and K. Ramchandran, "Network Coding for Distributed Storage Systems," *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4539-4551, Sept. 2010. [Article \(CrossRef Link\)](#)
- [4] K. V. Rashmi, N. B. Shah and P. V. Kumar, "Optimal Exact-Regenerating Codes for Distributed Storage at the MSR and MBR Points via a Product-Matrix Construction," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 5227-5239, Aug. 2011. [Article \(CrossRef Link\)](#)
- [5] T. Zhou, H. Li, B. Zhu, Y. Zhang, H. Hou and J. Chen, "STORE: Data recovery with approximate minimum network bandwidth and disk I/O in distributed storage systems," in *Proc. of 2014 IEEE International Conference on Big Data (Big Data)*, Washington, DC, USA, pp. 33-38, 2014. [Article \(CrossRef Link\)](#)
- [6] D. S. Papailiopoulos and A. G. Dimakis, "Locally Repairable Codes," *IEEE Transactions on Information Theory*, vol. 60, no. 10, pp. 5843-5855, Oct. 2014. [Article \(CrossRef Link\)](#)
- [7] D. S. Papailiopoulos, J. Luo, A. G. Dimakis, C. Huang and J. Li, "Simple regenerating codes: Network coding for cloud storage," in *Proc. of 2012 Proceedings IEEE INFOCOM*, Orlando, FL, USA, pp. 2801-2805, 2012. [Article \(CrossRef Link\)](#)
- [8] S. El Rouayheb and K. Ramchandran, "Fractional repetition codes for repair in distributed storage systems," in *Proc. of 2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Monticello, IL, USA, pp. 1510-1517, 2010. [Article \(CrossRef Link\)](#)
- [9] B. Zhu and H. Li, "Adaptive Fractional Repetition Codes for Dynamic Storage Systems," *IEEE Communications Letters*, vol. 19, no. 12, pp. 2078-2081, Dec. 2015. [Article \(CrossRef Link\)](#)
- [10] B. K. Rai, V. Dhoorjati, L. Saini and A. K. Jha, "On adaptive distributed storage systems," in *Proc. of 2015 IEEE International Symposium on Information Theory (ISIT)*, Hong Kong, China, pp. 1482-1486, 2015. [Article \(CrossRef Link\)](#)
- [11] O. Olmez and A. Ramamoorthy, "Repairable replication-based storage systems using resolvable designs," in *Proc. of 2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Monticello, IL, USA, pp. 1174-1181, 2012. [Article \(CrossRef Link\)](#)
- [12] O. Olmez and A. Ramamoorthy, "Constructions of fractional repetition codes from combinatorial designs," in *Proc. of 2013 Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, pp. 647-651, 2013. [Article \(CrossRef Link\)](#)
- [13] O. Olmez and A. Ramamoorthy, "Fractional Repetition Codes With Flexible Repair From Combinatorial Designs," *IEEE Transactions on Information Theory*, vol. 62, no. 4, pp. 1565-1591, Apr. 2016. [Article \(CrossRef Link\)](#)
- [14] Y. S. Su, "Constructions of Fractional Repetition Codes with Flexible Per-Node Storage and Repetition Degree," in *Proc. of GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, Singapore, pp. 1-6, 2017. [Article \(CrossRef Link\)](#)
- [15] R. Diestel, "The Basics," in *Graph Theory*, 5th ed. Springer-Verlag, Heidelberg, GER, 2016, ch. 1, sec. 10, pp. 27-28. [Online]. Available: <https://www.diestel-graph-theory.com/eBook.html>
- [16] J. Kim and H. Song, "Hypergraph-Based Binary Locally Repairable Codes With Availability," *IEEE Communications Letters*, vol. 21, no. 11, pp. 2332-2335, Nov. 2017. [Article \(CrossRef Link\)](#)
- [17] B. Zhu, "A Study on Universally Good Fractional Repetition Codes," *IEEE Communications Letters*, vol. 22, no. 5, pp. 890-893, May 2018. [Article \(CrossRef Link\)](#)
- [18] K. Gopal, M. K. Gupta, "Bounds on generalized FR codes using hypergraphs," *Journal of Applied Mathematics and Computing*, vol. 65, no. 1, pp. 771-792, Feb. 2021. [Article \(CrossRef Link\)](#)

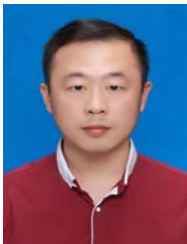
- [19] J. Wang, K. Shen, X. Liu and C. Yu, "Construction of Binary Locally Repairable Codes With Optimal Distance and Code Rate," *IEEE Communications Letters*, vol. 25, no. 7, pp. 2109-2113, Jul. 2021. [Article \(CrossRef Link\)](#)
- [20] B. Zhu, S. Zhang and W. Wang, "Expandable Fractional Repetition Codes for Distributed Storage Systems," in *Proc. of 2021 IEEE Information Theory Workshop (ITW)*, Kanazawa, Japan, pp. 1-5, 2021. [Article \(CrossRef Link\)](#)



Tiantian Wang is currently a teaching assistant with the Naval Aviation University. She received her M.S. degree in the School of Information Engineering from Chang'an University, Shaanxi, China, in 2019. Her research interests are in the area of network coding, distributed storage and regenerating codes.



Jing Wang is currently a professor with the School of Information Engineering, Chang'an University, China. She received her B.S., M.S. and Ph.D. degrees in the School of Communication Engineering from Xidian University, Shaanxi, China, in 2004, 2005, and 2009, respectively. Her research interests are in the area of network coding, distributed storage and regenerating codes.



Haipeng Wang received the Ph.D. degree from Naval Aviation University, in 2012, where he is currently a professor. His research interests include the general area of intelligent perception and fusion, and big data technology and application. He also serves as a Reviewer for several distinguished journals, including IET RSN and IEEE AES.



Jie Meng is currently a teaching assistant with the Naval Aviation University. She received her M.S. degree from China University of Petroleum (East China), Shandong, China, in 2019. Her research interests are in the area of Machine learning, pattern recognition.



Chunlei Yu received his M.S. in the School of Information Engineering from Chang'an University, Shaanxi, China, in 2020. His research interests include distributed storage, regenerating codes and locally repairable codes.



Shuxia Wang is currently a teaching assistant with the Shanxi Jinzhong Institute of Technology. She received her M.S. degree in the School of Information Engineering from Chang'an University, Shaanxi, China, in 2019. Her research interests include distributed storage, regenerating codes and locally repairable codes.