Special focus on distributed storage coding*

Due to their low cost and feasibility, distributed storage systems (DSSs) have been widely applied in large data centers, peer-to-peer storage systems such as OceanStore, Total Recall, and DHash++, and for storage in wireless networks. However, as the storage scale dramatically increases, reliability becomes a critical issue in applications involving the DSSs. The most efficient way to guarantee reliability is to employ erasure coding. However, it brings many new challenges that need to be addressed in terms of scalable deployment, security, efficiency, new devices, and so on.

This special focus of *SCIENCE CHINA Information Sciences* is devoted to covering the recent progress in coding/decoding methods and their security. We have accepted seven invited papers with novel results and techniques in these fields as follows.

In a DSS, an important issue is the efficient repair of a failed node, where data are dispersed across other nodes. The survey paper "Erasure coding for distributed storage: an overview" provides an overview of the advances in erasure coding with efficient repairs that have taken place in the past decade; in particular, it provides an overview of regenerating codes and locally recoverable codes, as well as novel ways to repair the ubiquitous Reed-Solomon code.

Among all erasure codes, the maximum distance separable (MDS) code is very desirable for practical systems in terms of redundancy and reliability. The paper "A class of binary MDS array codes with asymptotically weak-optimal repair" proposes a class of binary MDS array codes with $k \geqslant 3$ information columns and $r \geqslant 2$ parity columns that asymptotically achieve the weak-optimal repair of an information column with $k + 1$ helper columns.

In the heterogeneous DSSs, nodes are divided into different clusters (racks). To reduce the communication cost, downloading more data from intra-cluster nodes is preferred. The paper "Storage and repair bandwidth tradeoff for distributed storage systems with clusters and separate nodes" investigates the optimal tradeoff for the cluster DSSs under general storage/repair parameters. Further, a regenerating code construction strategy is proposed to achieve the optimal tradeoff curve for the cluster DSSs with specific parameters.

In some practical scenarios, such as in large-scale storage systems with correlated failure patterns, multiple failures are common. The paper "Code constructions for multi-node exact repair in distributed storage" studies the problem of the centralized exact repair of multiple failures in distributed storage. In particular, the authors present constructions achieving a new set of interior points under exact repair and prove the optimality of one point under the functional repair.

Flash memory is widely used in storage systems because of its high density and relatively low cost. The rank-modulation scheme can efficiently represent data in flash memory. The paper "Gray codes over certain run-length sequences for local rank modulation" presents Gray codes over certain run-length sequences in the local rank-modulation scheme that can overcome the drawback of the many distinct charge levels required in the rank-modulation scheme. Furthermore, the proposed codes are demonstrated to have an asymptotically optimal rate.

Private information retrieval (PIR) arises from the secure retrieval of information from public storage systems. The paper "On sub-packetization and access number of capacity-achieving PIR schemes for

MDS coded non-colluding servers" establishes lower bounds on the sub-packetization and access numbers of coded PIR schemes. Further, a general linear capacity-achieving coded PIR scheme is presented that attains the bounds simultaneously.

Diversity coding, node repair, and security are three basic ingredients of modern DSSs. The paper "New results on multilevel diversity coding with secure regeneration" characterizes the minimum-bandwidth-regeneration (MBR) point of the achievable normalized storage-capacity repair-bandwidth tradeoff region under the assumption that the eavesdropper can access the repaired data on some compromised storage nodes, but only the data on the remaining compromised nodes, as long as the number of the former compromised nodes is no greater than that of the latter compromised nodes.

Finally, we would like to express our sincere appreciation to all the authors for their High-quality manuscripts, as well as all the reviewers for the timely constructive comments. We must thank Dr. Jing Feng, the Executive Editor at the *SCIENCE CHINA Information Sciences* office, for the tremendous editorial support during the whole process of this special focus.

Guest Editors:
Xiaohu TANG
Southwest Jiaotong University, China
Shu-Tao XIA
Tsinghua University, China
Chao TIAN
Texas A&M University, USA
Qin HUANG
Beihang University, China
Xiang-Gen XIA
University of Delaware, USA