INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
SHORT ABSTRACT OF THESIS

Name of the Student	:	Ujwal Deep Kadiyam
Roll Number	:	136102027
Programme of Study	:	Ph.D.

Thesis Title: Codes with availability and multiple local erasure correction for distributed data storage.

Name of Thesis Supervisor(s)	:	Dr. Smarajit Das
Thesis Submitted to the Department/ Center	:	Electronics and Electrical engineering
Date of completion of Thesis Viva-Voce Exam	:	28-4-2021
Key words for description of Thesis Work	:	Locally recoverable codes, locality, availability, minimum Hamming distance, Distributed storage systems, multiple local erasure correction.

SHORT ABSTRACT

In any data storage system, the system should provide the data even in the presence of disk failures. Coding theory is used to provide these features in Distributed Data storage systems. Locally Recoverable codes (LRCs) are used to recover data from a failed node in distributed data storage systems. LRCs with availability are proposed to provide accessibility to a disk by multiple users at the same time. In a locally recoverable code (LRC), any code symbol can be recovered by accessing at most r other symbols (called a recovery set). In an LRC with *availability ' t'*, any information symbol has t disjoint recovery sets.

As the first result in this thesis, we consider a new class of codes with availability, where each disjoint recovery set for any information symbol is allowed to have different locality and each disjoint recovery set is allowed to be protected by local code of different minimum Hamming distance. An upper-bound on the minimum Hamming distance of these codes is derived. A family of systematic codes with information availability are constructed achieving the bound with equality. We show that these optimal codes provide availability for any information symbol even when some symbols are erased from its disjoint recovery sets.

As the second result in this thesis, we consider a new class of codes with availability. These codes provide availability for a subset of code symbols instead of individual code symbols. These are referred to as codes with multi-symbol availability. For this code, an upper-bound on minimum Hamming distance is derived and optimal codes with information multi-symbol availability have been constructed attaining the bounds with equality.

As the third result in this thesis, we define codes with availability, where the disjoint recovery sets for a code symbol are not fixed. One can choose t subsets, each consisting of any r code symbols to provide availability for a code symbol. We derive an upper bound on the minimum Hamming distance of this code and provide optimal code constructions. We show the applications of these codes in blockchain systems.