

**CHENNAI – PONDICHERRY**

**ENHANCED SECURE THRESHOLDED DATA DEDUPLICATION SCHEME FOR CLOUD STORAGE**

**Abstract:**

As more corporate and private users outsource their data to cloud storage, recent data breach incidents make end-to-end encryption increasingly desirable. Unfortunately, semantically secure encryption renders various cost-effective storage optimization techniques, such as data deduplication, ineffective. On this ground Stanek et al. [1] introduced the concept of “data popularity” arguing that data known/owned by many users do not require as strong protection as unpopular data; based on this, Stanek et al. presented an encryption scheme, where the initially semantically secure ciphertext of a file is transparently downgraded to a convergent ciphertext that allows for deduplication as soon as the file becomes popular. In this paper we propose an enhanced version of the original scheme. Focusing on practicality, we modify the original scheme to improve its efficiency and emphasize clear functionality. We analyze the efficiency based on popularity properties of real datasets and provide a detailed performance evaluation, including comparison to alternative schemes in real-like settings. Importantly, the new scheme moves the handling of sensitive decryption shares and popularity state information out of the cloud storage, allowing for improved security notion, simpler security proofs and easier adoption. We show that the new scheme is secure under the Symmetric External Diffie-Hellman assumption in the random oracle model.