A Scalable Data Chunk Similaritybased

Compression Approach for Efficient

Big Sensing Data Processing on Cloud

**ABSTRACT:** 

Big sensing data is prevalent in both industry and scientific research

applications where the data is generated with high volume and velocity.

Cloud computing provides a promising platform for big sensing data

processing and storage as it provides a flexible stack of massive computing,

storage, and software services in a scalable manner. Current big sensing data

processing on Cloud have adopted some data compression techniques.

However, due to the high volume and velocity of big sensing data, traditional

data compression techniques lack sufficient efficiency and scalability for data

processing. Based on specific on-Cloud data compression requirements, we

propose a novel scalable data compression approach based on calculating

similarity among the partitioned data chunks. Instead of compressing basic

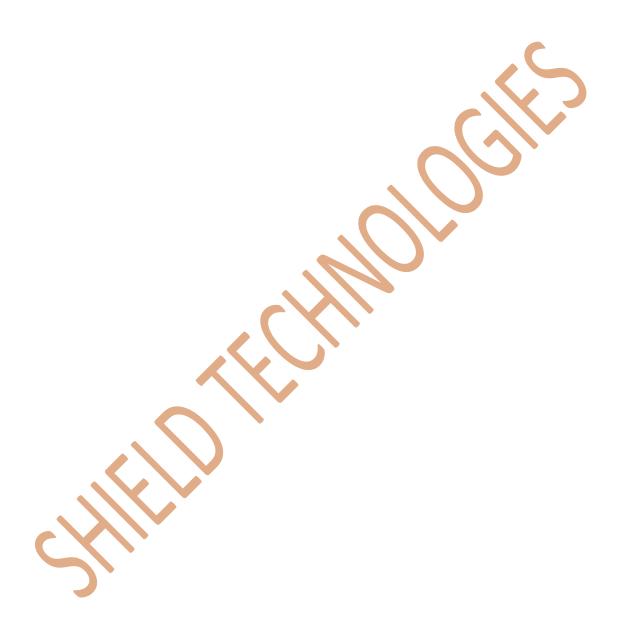
data units, the compression will be conducted over partitioned data chunks.

To restore original data sets, some restoration functions and predictions will

SHIELD TECHNOLOGIES,

be designed. MapReduce is used for algorithm implementation to achieve extra scalability on Cloud.

selective encryption strategy within the required execution time requirements.



2232, 3<sup>RD</sup> FLOOR, 16<sup>TH</sup> B CROSS, YELAHANKA NEW TOWN, BANGALORE-64

Mail us: <a href="mailto:shieldtechnoblr@gmail.com">shieldtechno.com</a> / <a href="mailto:manager@shieldtechno.com">manager@shieldtechno.com</a>

Contact: 9972364704 / 8073744810