

**CHENNAI – PONDICHERRY**

**Towards Privacy Preserving Publishing of Set-valued Data on Hybrid Cloud**

**Abstract:**

Storage as a service has become an important paradigm in cloud computing for its great flexibility and economic savings. However, the development is hampered by data privacy concerns: data owners no longer physically possess the storage of their data. In this work, we study the issue of privacy-preserving set-valued data publishing. Existing data privacy-preserving techniques (such as encryption, suppression, generalization) are not applicable in many real scenes, since they would incur large overhead for data query or high information loss. Motivated by this observation, we present a suite of new techniques that make privacy-aware set-valued data publishing feasible on hybrid cloud. On data publishing phase, we propose a data partition technique, named extended quasi-identifier-partitioning (EQI-partitioning), which disassociates record terms that participate in identifying combinations. This way the cloud server cannot associate with high probability a record with rare term combinations. We prove the privacy guarantee of our mechanism. On data querying phase, we adopt interactive differential privacy strategy to resist privacy breaches from statistical queries. We finally evaluate its performance using real-life data sets on our cloud test-bed. Our extensive experiments demonstrate the validity and practicality of the proposed scheme.

**Existing System:**

The existing data publishing technologies, data querying is sucha punchbag. As there are a thousand Hamlets in a thousand people's eyes, differential data privacy protec-tion should be provided for satisfying the various re-quirements on users withdifferent roles. User roles reflect the commercial agreements between data owner and au-thorized users. Similar to fine-grained access control, data owner should provide different granular results for au-thorized users with different roles, even facing the same data set for the same query.

**Proposed System:**

To the best of our knowledge,this is the first study that formalizes the problem of privacy-preserving set-valued data publication over hybrid cloud, and provides a complete system framework. Our design proposes a novel data partition mechanism that splits EQI into dif-ferent chunks, and it ensures that private information will not be exposed to the public cloud. Moreover, we employ interactive differential privacy into the proposed frame-work, which provides strong privacy guarantees.

We have implemented our scheme and achieved limited information loss by evaluating it over our cloud testbed on real datasets. Furthermore, our extensive ex-periments further demonstrate the validity and practicali-ty of our scheme.