# EFFICIENCT QUERY SERVICES AND SECURE FILE RETRIEVAL IN BUSINESS CLOUDS

**Vanajakshi Devi.K[1], Kalpana.D[2], Shahul Ameed.P[3]**

[1]Vemu Institute of Technology
Pakala, Chittoor, India

[2]Vemu Institute of Technology, Department of CSE,
Pakala, Chittoor, India

[3]Yogananda Institute of Technology & Science, JNTUA,
Mohan Reddy Nagar, Tirupati, India

**Abstract:** ***Cloud computing loom as a hottest tendency in field of information technology and composite networking. Due to less prudent quota resources in a cost-efficient cloud habitat, an end user can countenance a certain scope of delay while fetching knowledge from the cloud to curtail costs. In this paper, we spotlighted on two essential issues in such a habitat: search confidentiality and competence. We first come across an exclusive searching scheme that was originally proposed by Ostrovsky. Their exclusive keyword based regeneration blueprint allows a user to reclaim files of interest from un-trusted servers in emission of any knowledge. The main drawback is because of altering of all the queries from different users, it will cause an excessive querying overhead aroused on the cloud and thus goes against the initial objective of cost adaptability. In this paper, we commenced three adequate information rehabilitation for ranked query (EIRQ) schemes to reduce querying overhead aroused on the cloud. In EIRQ, queries are classified into multiple ranks, where a greater ranked query can recover a greater percentage of paired files on user demand. A user can reclaim files on demand by choosing queries of distant ranks. This aspect is useful when the user only needs a small fragment of files from enormous number of matched files. This entity introduces retrieval of files with low bandwidth and low computational and communication cost. Under different parameter settings, extensive evaluations have been conducted on both discrete models and on a real cloud habitat, in order to audit the effectiveness of our blueprints.***

**Keywords:** ADL, rank matrix, Ostrovsky, EIRQ, user privacy



## 1. Introduction

Cloud computing enables convenient, pervasive, on-demand network access to a mutual pool of configurable computing resources that can be promptly maintained and liberated with nominal management intention [3]. As clouds are cost active, flexible and extensible, many organizations are looking forward to dispose their knowledge for sharing. In clouds, files are analyzed by keywords [6] with a query and reclaim the files that are implicated. The objective for the cost efficient clouds is to cost of CPU utilization, cost of network bandwidth usage, increase privacy of users. In this scenario, the cohesion of user will be major issue. User privacy [1] can be assorted into search privacy and access privacy. Exploration privacy accord with the user's exploration and access privacy deals with the fields that are to be rebounded. Ostrovosky scheme [4] is antique as it has to action all the keywords in barring no one file. This process edges to substantial query overhead from distinct users. An avant-garde solution would be to make a rank matrix that intensify the user privacy than the erstwhile methods. EIRQ scheme address the issues of privacy, aggregation, computational cost and bandwidth wastage.

## 2. OVERVIEW

Private searching was proposed by Ostrovsky allows user to fetch files of interest from an untrusted server without disclosing any information. Contrarily, the cloud will learn that certain files, without processing, are of no interest to the user. Commercial clouds follow a pay-as-you-go model [6], where the customer is publicized for different transactions such as bandwidth, CPU eternity, and so on. To make private searching pertinent in a cloud habitat, our prior work designed a cooperate private searching protocol (COPS) [7], where a surrogate server, called the aggregation and distribution layer (ADL), is imported between the users and the cloud. Our intent is to conserve user privacy through differential query services by Aggregation and distributions Layer. The ADL positioned inside an organization has two principal functionalities: aggregating user queries and distributing search results. Under the ADL, the enumeration cost fetched on the cloud can be largely curtailed, since the cloud only needs to carry out a blended query once, no matter how many users are executing queries. Nevertheless, the transmission cost fetched on the cloud will also be abbreviated, since files shared by the users need to be recompensed only once. Most substantially, by using a sequence of secure functions, COPS can preserve user privacy from the ADL, the cloud, and other users. The issues with existing scheme has an eminent computational

cost [1], since it constrains the cloud to approach the query on every file in a collection. It will promptly become a performance obstruction when the cloud needs to process thousands of queries over a mass of hundreds of thousands of files [2]. That is the reason we switch our impulse of research towards differential query services with ADL in order to reduce computational cost, low bandwidth usage.

## 3. System Architecture

### 3.1 Model

In Co-operate searching protocol, ADL acts as a proxy server [8] and as a moderator between user and cloud. The three entities in the system model are varied users, ADL and cloud. When a user consigns queries to cloud it is first reached to ADL, where all the queries from the users are aggregated and dispatched to the cloud as a compound query [7]. These queries are refined by the cloud and dispatches the file buffer to ADL. ADL forthwith takes the liability of dispensing the files to its equivalent users. Thus the bandwidth usage is minimized.

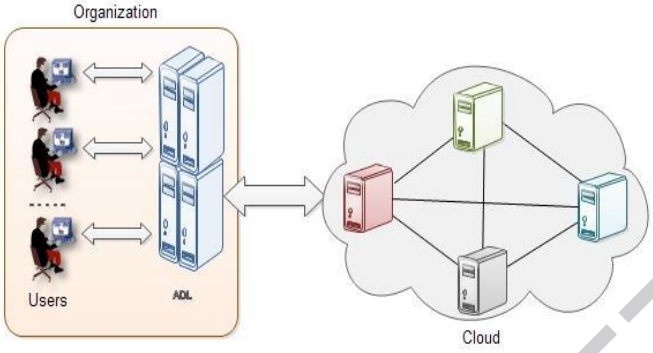


**Figure 1:** System Model

## 4. Proposed System

In this paper, we initiate a significant concept, differential query services, to COPS, where the users are authorized to intimately judge how many matched files will be refunded. This is impelled by the phenomenon that beneath certain cases, there are a pile of files corresponding a user's query, but the user is concerned in only a definite percentage of matched files. In the Ostrovsky scheme, the cloud will have to recompense 4,000 files. In the COPS scheme, the cloud will have to recompense 2,000 files. In our scheme, the cloud only needs to return 400 files. Therefore, by authorizing the users to retrieve matched files on demand, the bandwidth squandered in the cloud can be largely lessened. Efficient information retrieval for Ranked Query (EIRQ), in which queries are sorted into varied ranks, where an eminent percentage of similar files can retrieved by higher ranked query. The simple notion of EIRQ is to frame a privacy-preserving mask matrix which makes the cloud to screens certain percentage of similar files before refunding to the ADL for aggregation. This is not an insignificant work, since the cloud needs to precisely filter out files in accordance to the rank of queries beyond knowing nothing about user privacy. By permitting the users to retrieve matched files on requisition, the bandwidth expended in the cloud can be largely lowered. We afford two resolutions to adapt associated parameters; solitary based on the Ostrovsky scheme, and another is based on Bloom filters.

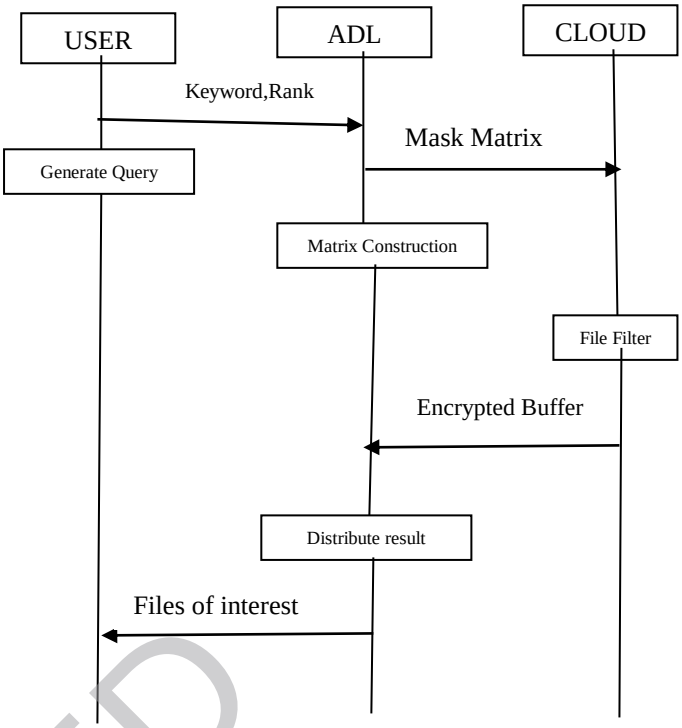


**Figure 2:** The EIRQ Scheme

## 5. Results & Discussion

The EIRQ schemes peruses the contrast of the computational/communication price, file durability rates fetched on the clouds with distinct knowledge retrieval schemes. File durability rates give the mobility of retrieving the in demand file for a user using Ostrovsky scheme and our model. Here the queries are categorized into 0~4 ranks Rank 0 to Rank 4. These ranks retrieve files of the user fascinated with the percentages of 100%, 75%, 50%, 25%, 0%.

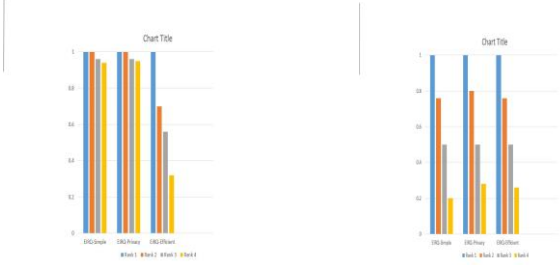

**Figure 3:** File Survival Rate under Bloom Filter Setting

## 6. Future Work

The differential queries in EIRQ scheme provides the required services by using cloud infrastructure. In this paper, we intended to enhance the experience to new customers and global markets by helping the partners augment their cloud contribution. EIRQ protocol is the efficient than Ostrovosky, COPS scheme as it addresses issues of privacy, aggregation and network bandwidth usage. The EIRQ can be further enhanced in future in certain aspects like:

1. This kind of ranking i.e. ranking of files depends upon highest rank of queries leaves few bottlenecks, a sophisticated ranking system can be developed by providing attributes to each file.

2. In EIRQ algorithm, the privacy preserving mask matrix has a row for each keyword in organization directory which threats the scalability of organization that has dictionaries with thousands of keywords. Still a reliable version of this algorithm can be proposed to compress the size of matrix.

## 7. Conclusion

In business organizations, cloud provides many advantages but few cons are user privacy, security and efficiency. By using query services that provides a bridge between empty set and the desired result set for the users. Here we use ranked query services so as to make the file retrieval easy and cost efficient with less bandwidth usage along with computational cost. The EIRQ methods used provides multi keywords search along with rank for the queries so as to make the file access easy with enhanced user privacy. EIRQ scheme based on the ADL, advances differential query services that conserves user's privacy, a user can redeem distinct percentage of matched files relying against the queries of distinct ranks. Private searching procedures are made more fitting to a cost efficient cloud habitat by using EIRQ schemes. Moreover to that, it extremely abbreviates the communication cost annexed on the cloud. Yet, here we vindicate the rank of each file by the highest rank of queries it matches and user retrieves definite percentage of files.

## 8. References

[1] G.Satya Suneetha, S.V.Ramana Murthy, T.S.V.V.S.Savithri," EIRQ Methods to Provide a Cost-Efficient Solution for Private Searching in Cloud Computing", In IJCST, 2014.

[2] Qin Liu, Chiu C. Tan, Jie Wu, and Guojun Wang, "Efficient Information Retrieval for Ranked Queries in Cost Effective Cloud Environment", IEEE INFOCOM.

[3] R. Ostrovsky and W. Skeith III, "Private searching on streaming data," in Proc. Of ACM CRYPTO,2005.

[4] Qin Liu, Chiu C. Tan, Jie Wu, and Guojun Wang, "Towards Differential Query Services in cost-Efficient clouds" IEEE Transactions on Parallel and Distributed Systems,Volume:25,Issue:6,Issue Date :June.2014

[5] Dr. G. Prakash Babu, B. Thirupthaiah," Cost Efficient Query Services through Aggregation and Distribution Layer", In IJAREEIE, July 2014.

[6] Cong Wang, Ning Cao, Kui Ren, and Wenjing Lou, "Enabling Secure and Efficient Ranked Keyword Search over Outsourced Cloud Data", IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, SYSTEMS, VOL. 23, NO. 8.

[7] Qin Liu, Chiu C. Tan, Jie Wu, and Guojun Wang, "Towards Differential Query Services in cost-Efficient clouds" IEEE Transactions on Parallel and Distributed Systems,Volume:25,Issue:6,Issue Date :June.2014

[8] Hore, S. Mehrotra, and G. Tsudik, "A privacy-preserving index for range queries," in Proceedings of Very Large Databases Conference (VLDB), 2004.

[9]C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, "Secure ranked keyword search over encrypted cloud data," in Proceedings of IEEE International Conference on Distributed Computing Systems (ICDCS), 2010.

## Author Profile

**VANAJAKSHI DEVI.K,** pursuing her Master Degree in Vemu Institute of Technology, pakala, Chittoor. She has membership in MISTE.

**KALPANA.D**, received her MASTER's Degree and working as Assistant Professor in Vemu Institute of Technology. She has attended several Conferences.

**P.SHAHUL AMEED,** pursuing Bachelor's Degree in Yogananda Institute of Technology & Science, Tirupati. He also has attended several National & International conferences.