A Range Query Algorithm to Process KNN Queries in Cloud Computing

[1] S. Selva Kanmani, [2] K. Vinupriya, [3] M. Yamuna,
[4]R. Sudheendra Deepika
[1] Head of The Department, Dept. of IT
[2][3][4] B.Tech Final Year Students (IT),
Velammal Institute of Technology, Panchetti,
Tamilnadu, India.

Abstract:- In Public Cloud environment, security and data confidentiality is the major problem facing by all the data controller. The service user can take the service from the cloud by getting authorization from the service provider and he can only pay for the service by using the server, for that the service provider lose the control so there may have chances of leaking the information. For that reason the data controller does not want the data to shift to the Cloud. There may have only chance of storing the data is providing the privacy gurantee to the Cloud. The requirement to building privacy is based on CPEL criteria which is confidentiality, privacy, efficiency, low in-house processing cost. By satisfying these requirement will increase the difficulty to store data in the Cloud. In order to eliminate this problem, We are using Random Space Perturbation method for providing the security and efficiency for processing the data. This method is used to building a practical query services in the Cloud. This approach will balance all the requirements by using range query and KNN query services.It provides multifaceted distances , which allows actual listing techniques to increase distances processing.

Keywords:- Query privacy, range query services , KNN query services, CPELcriteria, RASP.

_____*****_____

I. INTRODUCTION

In Cloud environment conversing query services is one of the solution for extending the ability of the system processing and cost-saving. But most of the approaches focusing only on the confidentiality but the in-house processing cost will be increased but this is no more use. This can already addressed by M. armburst et al, this will increase the in-house infrastructures [1] however, access provider fail to control the data in the Cloud, so security have become problem in public cloud. some of the service provider can leaking the information to authorized user based on the Data as a service. While some of the actual system are needed to conserve the data security and confidentiality gurantee then only we can get the ful benefits of the cloud. But it is very difficult to provide such privacy in cloud infrastructures. Therefore, there is an some relationship among confidentiality, query privacy cost saving in Cloud.

II. EXISTING SYSTEM

The requirements are used for building the effective query services in cloud by using these requirements are confidentiality, privacy, efficiency, low in house cost. To fulfilling these requirement will promptly increase the difficulty by building the practical query services in the Cloud. Some of the approaches have been developed to solve these problem. But they do not satisfying these

requirements. In OPES scheme in [2] are vulnerability to attacks because they do not preserve the data approximately.

III. OTHER RELATED WORK

R. Agrawal et al[2] focuses on the technology Encryption is used for protecting sensitive data. But once encrypted also easily accessed from the exact matches. In this order preserving schemes can allow any comparison operator that can be used in encrypted data. some of the aggregation function are used in comparison operator and similarly GROUPBY and ORDER BY can also be used. In these scheme it preserves the order and the processing over data encryption are also approximate. This will contain neither false positives nor miss any tuples. This eill exactly differentiates it other scheme. This scheme handles updates gracefully. The OPES will reduce the vulnerability to the attacks. However the hacker can get access to the information retrival but he cannot able to view the data. The overhead on query processing will be reduced but encryption key is large and it occupies more space.

Zhengli Huang et al[3]describes that Randomization has the convienient method for data storing. The data gathering and data searching have increases both in space and difficulty. The randomization technique does not provide privacy. However, it has to provide clear reason which causing security problem, what are the reason for causing such security problem and how to give protection to security issues, and which data needs higher security. It is cleared in

1425

this approach by anatomization. The research of this method (ppdm) [3] aimed at filling the space merged data storing and data privacy. It searches on many possible ways in social issues in general. They have analysed and taken many experimental analysis on the connection between data combining and the amount of secure information that can be providing security based in this proposed data techniques. To improve security, the proposed a modified anatomization scheme, in which we let the combimation of random noises similar to the actual data.

W. K. K et al[4] describes that the service providers convert their infrastructures into public cloud environment. Some of the traditional encryption method focus on providing unbreakable protection even though they do not holds up the queries structures on the inculpation data. It focus on the information retrieval on inculpation and they develop a various new inculpation method which provides secure scalar product. The purpose of the schemes is to protect the data attacks in different level, at a different overhanging cost. It indicating the security warning in Cloud environment. This method converting the plain text to cipher text using encryption and decryption method only authorized parties can perform decryption.

IV.   PROPOSED SYSTEM

The proposed Random space perturbation(RASP) approach is used for building efficient query service and providing confidential gurantee. RASP is combining the range query and KNN-R algorithm for providing query privacy to the data. The KNN-R algorithm which deals with the inner and outer range. The data stored by the data controller in the cloud will securely transformed into an imaginary solution based on the RASP data perturbation method. The RASP method which combines multidimensional data sets. The data controller is the responsibility for storing the data, depends on that only RASP method which securely transform the data into polyhedral.
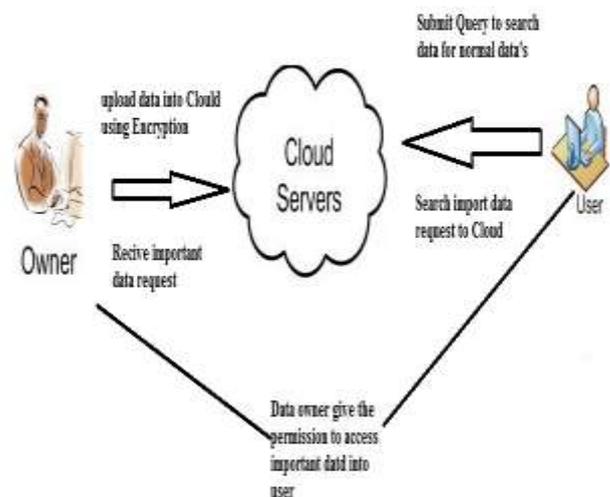
The KNN-R algorithm processing the range queries. In the proposed approach has some unique contribution they are:

- The RASP is the method to combine many schemes such as randomization, encryption, multi dimensional expansion to provide privacy guarantee.
- The RASP method preserves multidimensional range and allows indexing to speed up the processing.
- One of the advantages of the nproposed system is that it reduces the high task at hand due to the less cost involved and increased meticulousness in the result.

- The above characteristics help in giving the required Cloud-based solutions in the most efficient manner.

System Architecture

In the Cloud environment, some of them are Amazon EC2[1],icloud. Amazon ec2 is a web service that resizable compute capacity in the cloud. In this architecture it purpose is to expand information retrival server to the hybrid cloud or using private cloud. This will improves system structures and ability. The data controller stores encrypted information to the cloud. the authorized user can search data in the cloud by submitting the queries to find some information or any hacking purpose. The service provider who may hack the information but they were curious so it may have leaking of secured information. The RASP imaginary information will be used to construct listing process to increase query processing.

V.   KNN-R THEOREM

The RASP method does not provide distances so it combines the knn query that is k nearest neighbor is used to find the ranges. Range query forms the basis of constructing the processing theorem[1]. The usage of pointer indicator i. e, index in this processing provides an efficient and fast way to deal with processing of queries. It shows the query result based on the inner range query.

VI.   ILLUSTRATION KNN-R Algorithm:
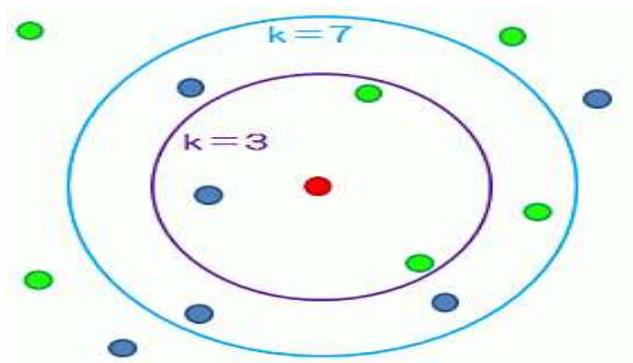
This illustration shows how the confidential range and general queries range are processing based on the query point. Based on the illustration only the authorized user query will be processing and it finds the inner range based on the outer range only it finds the outer range query result.

VII.   CONCLUSION

The basic need to understand the advantages of cloud computing is to have less work-pressure in the system. This will provide more efficient and secure query services in the public Cloud infrastructures.

REFERENCES

[1] Exploiting RASP Data Perturbation for Building Confidential and Query Services in the Cloud **S Priyanga, MS Lalitha - ijrste.org**

[2] R.Agarwal, J.Kiernan,R.Srikant, and Y.Xu, " Order preserving encrypted for numeric data" proc.ACM SIGMOD Int'l conf.management of data(SIGMOD),2004.

[3] Z.Huang,W.Du, and B.Chen, " Derving private information from randomized data " Proc.ACM SIGMOD Int'l conf. management of data(SIGMOD),2005.

[4] W.K.K,D.W.-L Cheung,B.Kao and N.Mamoulis, "Secure KNN computation on encrypted information retrivals," Proc.ACM SIGMOD Int'l conf. management of data(SIGMOD).

[5] S. Papadopoulos, S. Bakiras, and D. Papadias, "Nearest NeighborSearch with Strong Location Privacy," retrivalsConf. (VLDB), 2010

[6] M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R.K. Andy Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M.Zaharia, "Above the Clouds: A research of Cloud Computing,"technical report, Univ. of Berkerley, 2009.

[7] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data," Proc. IEEE INFOCOMM, 2011.

[8] C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, "Secure Ranked Keyword Search over Encrypted Cloud Data," Proc. IEEE Int'l Conf. Distributed Computing Systems (ICDCS), 2010.

[9] M.F. Mokbel, C. andG. Aref, "The New Casper: Query Processing for Location Services without Compromising Privacy," Proc. 32nd Int'l Conf. Very Large Information retrivals Conf. (VLDB), pp. 763-774, 2006.

[10] E. Shi, J. Bethencourt, T.-H.H. Chan, D. Song, and A. Perrig, "Multi-Dimensional Range Query over Encrypted Data,"