

# Spatial Data Querying System Using Security Approach and Accessing Model

Harsha Shrivastava<sup>1</sup>, Prof. Ratnesh Kumar Dubey<sup>2</sup>, Dr. Vineet Richhariya<sup>3</sup>

<sup>1</sup>M. Tech. Scholar, <sup>2</sup>Research Guide, <sup>3</sup>HOD

Department of Computer Science and Engineering, LNCT, Bhopal

**Abstract:** - Today a large amount of data is generated in each area. Spatial data is the importance of data where the usage, its access, storage is required in efficient manner. P2P system, communication makes use of these data with their efficient search and storage scheme. Security in access is a major concept where the requirement is to process, storage of data in secure format is required. In this paper our investigation in previous BestPeer++ Approach for large spatial data access with security parameter is monitored. A further extension with parallel approach and HECC (Hyper elliptic curve cryptography) security approach is taken in this paper which is compared with existing bestpeer++ approach over large dataset query system. It is observed that the proposed work outperform better than the existing algorithm in terms of execution time and security concern over the dataset TCP-H.

**Keywords-** Spatial data, query system, bestpeer++, HECC security.

## I. INTRODUCTION

Data mining from P2P community is a quite new area with little associated literature. Some researchers have developed numerous extraordinary procedures for computing simple operations (e.g. average, sum, max, execution time) on P2P networks. Query processing techniques for P2P systems are nevertheless inefficient: sending (atomic) queries to the suitable peers sincerely fails for queries which need records from multiple peers to be finished. While pretty some database techniques may be re-used, a P2P statistics control infrastructure poses additional demanding situations. Due to the fact of the dynamic nature of P2P networks, we are able to neither assume international information approximately statistics distribution, nor are static topologies and static question plans appropriate. In evaluation to traditional distributed database systems, we can't count on a whole schema but as alternative paintings with a disbursed allocation schema which

directs query processing tasks from one node the neighbors. In current day years, with the improvement of laptop network and database technology, allocated database is increasingly more widely used; with the increasing utility, statistics queries are increasingly more complicated, the performance requests are an increasing number of immoderate, so question processing is a key trouble of the allocated database gadget Query processing is a way for purchasing statistics from the database in a reliable way. The overall performance of the database device relies upon on the query processing techniques we utilized in database systems.

## II. RELATED WORK

BestPeer++ know-how administration platform. While average P2P neighborhood has now not been designed for group facets, the ultimate intention of BestPeer is to preserve the state-of art database approaches into P2P packages. In its early stage, BestPeer employs unstructured group and advantage retrieval system to examine columns of precise tables almost always. After defining the mapping points, queries will also be dispatched to particular nodes for processing. Best Peer introduces a sequence of systems of making enhancements to question effectively and affect fine to expand its suitability for organization neighborhood offerings. In certain, Best Peer provides robust dispensed search choices with a balanced tree structured overlay group (BATON [2]) and partial indexing scheme for reducing the index dimension. In addition, immoderate-excessive-high-quality-Peer develops adaptive emerge as a member of query processing [4] and dispensed online aggregation tactics to furnish effective question processing. Best Peer++ develops adaptive join query processing and distributed on-line aggregation strategies to provide efficient question processing .It is a cloud enabled understanding sharing platform designed for company community applications. By means of utilizing integrating cloud computing, database, and peer-to-peer (P2P) technologies, BestPeer++ achieves its question processing effectively and is a

promising method for corporation community functions, with the next distinctive points. BestPeer++ is more compatible with distributed entry manipulate, a few types of indexes, and pay-as-you-go query processing for supplying elastic knowledge sharing offerings within the cloud. The appliance add-ons of BestPeer++ are separated into two sections: core and adapter. The core entails all of the expertise sharing functionalities and is designed to be platforms neutral. The adapter involves one abstract adapter which defines the elastic infrastructure carrier interface and a set of concrete adapter add-ons which put into effect such an interface through APIs furnished through exact cloud provider companies (e.g., Amazon) and an Amazon cloud platform implemented by bestpeer+++.

### III. PROPOSED METHODOLOGY

#### 3.1 Problem Definition

As per the literature survey is performed with different techniques and different result from the algorithms were monitored such as Piazza[32] , HadoopDB[3], BestPeer++[1] and other different technique for query data processing on large amount of structured RDBMS available dataset our monitoring is performed. Upon verifying different scenario and the available technique different short comes with the Existing algorithm BestPeer++ which is taken as base for our research work. The following are the monitored points which identified as problem and further analyzed and performed further with enhancements. Previous technique such as HadoopDB utilized clustering technique where a high configuration is required to prepare and then a highly configured hardware requires processing the data. The technique HadoopDB work on the big data, which is having a dataset in structured manner with large data values and set for consideration under experiment. HadoopDB perform and required High RAM and Complex architecture for the experimental setup, thus a high computation time is required while analyzing the program output, while making it data transferrable and communicable with the large data and queries. BestPeer++ which is a technique based on peer not on Hadoop or big data processing framework perform better than HadoopDB but at the same time it uses some security consideration where it lowers the security constraints by using A PKI based

encryption system, which is not highly secure with the available market scenario, where different active attackers are available to steal the data and network. The existing technique computes high computation time, as the number of data processing an encryption rounds are more as compare to the proposed work given by our research.

#### 3.2 Proposed Methodology

As per our observation about the previous technique and their disadvantage in different terms and scenario's. Our work present a new approach which is highly secure and consumes low computational time and thus computational cost over the large number of structured available dataset. Our work propose a new algorithm SEBestPeer++ algorithm which is peer based high secure algorithm utilizes a highly proven symmetric key based encryption algorithm for the communication in between the Peer daemon process. The proposed algorithm utilizes Elliptic curve cryptography (ECC) algorithm for the communication message exchange in between the normal peer and bootstrap peer. The proposed algorithm is described below:

1. Listing and loading of the entire available normal peer in the network which is participating for the communication.
2. Creating an object of new normal peer.  
NormalPeernp=new NormalPeer ();
3. Perform communication in between normal peer and bootstrap peer using a secure algorithm ECC.
4. Perform key generation for ECC.
5. Perform encrypted data transmission over the bootstrap peer and normal active peer in the scenario.
6. Monitoring Metadata by bootstrap and normal peer.
7. Observing the execution time and thus it affect computational cost for the complete transmission.
8. Exit.

#### **Algorithm PsudoCode:**

*Secure bootstrap Daemon Algorithm (SEBestPeer++):*

*Input: Query Qi, Dataset tables DS.*

*Output: Communication process, Metadata,*

*Computation time.*

*Steps:*

```

normalpeer1=Inactive, normalpeer2=Inactive.
While (true) do {
    Peerlisting {peer1, peer2.....peerN};
    NewPeerCreationRequest ();
    NormalPeer1 normal1=new NormalPeer1();
    If(peercreation(>0)
    {
        Peercreationsuccess();
        Perform Communication Bootstrap to
NormalPeer using Q1;
        Apply ECC Encryption (Q1..Qn)
        {
            Perform authentication;
            Perform key generation using ECC points;
            Send Encrypted data to normalpeer;
        }
        Set status=Active; generate Metadata();
    }else
    {
        peerStatus=Inactive;
        generating metadata for peer request;
    }
    Bootstrap- remove activepeer();
}

```

**IV. IMPLEMENTATION DETAIL**

We put into impact the benchmark queries via ourselves when you recollect that the TPC-H queries are complex and time-consuming queries which aren't appropriate for benchmarking employer network packages. The TPC-H benchmark statistics set includes eight tables. We use the genuine TPC-H schema because of the truth the shared global schema. In Implementation the work is done in Java framework tool and result analysis here is the system tolerance detail I have applied some random click and observed my best analysis result. The proposed and existing technique is performed with the above query Q1, where the data is processed and following output results were monitored. In the above graph drawn x axis as computational time for the query processing for specified dataset and bar graph is printed using the chart library provided by the Microsoft and further analysis can easily performed thus the SEBestpeer++ outperform the best and low computational time with the same query number as the computation time is decreased due to low

encryption and decryption time using ECC encryption system.

Table 1: Statically analysis of obtained result

Technique Approach/ Query	Existing Bestpeer++ (ms)*	Proposed SEBestPeer++ (ms)*
Query 1	116695	99521
Query 2	18254	17023
Query 3	183015	162032
* Computation time in ms		

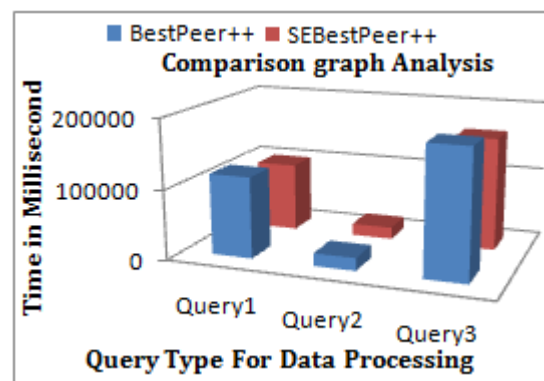


Figure 1 Different technique Comparison using Bar graph.

**V. CONCLUSION AND FUTURE WORK**

The work outperform by us based on the peer technique for the large data processing system , where a Bestpeer++ technique was introduced in the paper which is taken for consideration and further upon observing the problem formulation of security and its high computation time the motivation and objective is set to be more authentic. Thus the new proposed algorithm SEBestPeer++ is evolving by the work done by us. Which removes the drawback of PKI encryption technique and enhance the security using the power of symmetric key based encryption technique for the data transfer communication in between bootstrap and normal node in architecture? Also further the work can be performed on calculating more parameter such as throughput and efficiency.

**REFERENCES**

[1]. Gang Chen, Tianlei Hu, Dawei Jiang, Peng Lu, KianLee Tan, Hoang Tam Vo, and Sai Wu, "Extended BestPeer: A Peer-to-Peer Based

- Large-Scale Data Processing Platform” ,VOL. 26,NO. 6, JUNE 2014.
- [2]. H.V. Jagadish, B.C. Ooi, and Q.H. Vu, “BATON: A Balanced Tree Structure for Peer-to-Peer Networks,” Proc. 31st Int’l Conf. Very Large Data Bases (VLDB ’05), pp. 661-672, 2005.
- [3]. A. Abouzeid, K. Bajda-Pawlikowski, D.J. Abadi, A. Rasin, and A. Silberschatz, “HadoopDB: An Architectural Hybrid of Map Reduce and DBMS Technologies for Analytical Workloads,” Proc. VLDB Endowment, vol. 2, no. 1, pp. 922-933, 2009.
- [4]. S. Wu, Q.H. Vu, J. Li, and K.-L. Tan, “Adaptive Multi-Join Query Processing in PDBMS,” Proc. IEEE Int’l Conf. Data Eng. (ICDE ’09), pp. 1239-1242, 2009.
- [5]. Xiaochun Wang, Xiali Wang, and D. Mitchell Wilkes, Member, IEEE” A Divide-and-Conquer Approach for Minimum Spanning Tree-Based Clustering” ,IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 21, NO. 7, JULY 2009.
- [6]. Ashish Thusoo, Joydeep Sen Sarma, Namit Jain, Zheng Shao, Prasad Chakka, Suresh Anthony, Hao Liu, Pete Wyckoff and Raghotham Murthy” Hive A Warehousing Solution Over a Map Reduce Framework” Proc. VLDB Endowment, vol. 2, no. 2, pp. 1626-1629, 2009.
- [7]. Dan Boneh and Matt Franklin” Identity-Based Encryption from the Weil Pairing”, Appears in SIAM J. of Computing, Vol. 32, No. 3, pp. 586-615, 2003. An extended abstract of this paper appears in the Proceedings of Crypto 2001, volume 2139 of Lecture Notes in Computer Science, pages 213–229, Springer-Verlag, 2001.
- [8]. Ikshwansu Nautiyal and Madhu Sharma,” Encryption using Elliptic Curve Cryptography using Java as Implementation tool” International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 1, January 2014.
- [9]. Lipyeow Lim, Min Wang, Jeffrey Scott Vitter” SASH: A Self-Adaptive Histogram Set for Dynamically Changing Workloads”, Proceeding VLDB '03 Proceedings of the 29th international conference on Very large data bases - Volume 29,Pages 369-380.
- [10]. Anshu Chaturvedi, D.N. Goswami, Rishi Soni and Brijesh Kumar Chaurasia” Secure Multi-party Communication in Data-mining Applications”, International Journal of Database Theory and Application Vol.8, No.4 (2015), pp.299-306.
- [11]. Neha tripathi, Ganesan R” Data Security in Cloud Architecture Based on Diffie Hellman and Elliptical Curve Cryptography”, <https://eprint.iacr.org/2014/049.pdf>