ISSN:0975-9646

VD-kNN Query Processing over Untrusted Cloud

Utsav Raychaudhuri[#1], Chandni Singhal*[2], Neelima Shahi[#3], Shital Jadhav[#4], Mrs. Vaishali Kolhe[#5]

[#*]*Department Of Computer Engineering,*
D.Y. Patil College of Engineering, Akurdi, SPPU, India

***Abstract*--Interests of clients are distinctive. Gathering substantial amount of information for the clients is a very tough job. the data that is being sent to the clients are exceptionally delicate. This data belongs to the data owners who wouldn't want their data to be leaked or tampered by any third party. so this data is only sent to a few authenticated clients. the clients send their area of interest. the clients also need to think about closest points of interest in nearest neighbour framework. however the data owner does not have that much capability therefore he outsources his database to the cloud administration. here NN queries are being utilised on an untrusted cloud environment. however this carries a risk that the cloud administration may leak the data to a third party or the data may be given to an unauthenticated client. for this purpose mutable order preserving encoding is being used which is a request safeguarding encryption. to overcome all of these aforementioned problems the proposed system implements Vd-kNN algorithm to improve the security of the system. the system that is being developed guarantees that the privacy of the client as well as the data owner both are preserved.**

***Keywords*-- Location Privacy, Spatial Databases, Database Outsourcing, Mutable Order Preserving Encoding, Location Based Services**

I. Introduction

There are many area based administrations that are available for e.g.- maps, route and social networking. The constant redesigns of remote systems and cell phones leads to advancement that has made a difference. Many of these administrations furnish the client with sensitive data. Area-based administrations and area-ward inquiries have been utilizing the parcel of a hobby. The client needs to handle very surprising amount of sensitive information. Clients are given data about their points of interests. This data is with respect to their geolocation. The majority of the inquiries include area traits which is handled by the nearest–neighbour (NN) queries. The data owner needs to think about k POIs (e.g., eateries, exhibition halls, corner stores) that are closest to the client's present area (kNN). Gathering extensive measure of information for the client is exceptionally troublesome. At some point, data is exceptionally sensitive. So taking care of the data is exceptionally troublesome. Data stockpiling likewise is an enormous issue. So information proprietor does not make information open to all clients. It's permitted just to the paying clients. The client sends their present area guides and needs to the cloud administration, which sorts the closest POIs in NN. However information Owner does not have the that much stockpiling limit. So cloud administration is being utilized. Diverse sort Service like Google, Amazon are giving their data to the SaaS (Software as an Administration) business. Cloud gives power full stockpiling requiring little to no effort. But Cloud is not completely trusted. To permit questions preparing of NN inquiries in not a completely trusted environment while in the meantime ensuring each the POI and questioning clients' positions. This procedure is called mutable order preserving encoding (mOPE). It is a protected request-pholding encryption. An execution enhancement is taken under consideration for lessening the computational expense. A broad execution assessment of the system is being presented reasonably. Two routines are being presented. One of the routines is VD- kNN system for securing the NN inquiries. This works by handling scrambled queries. The strategy returns precise one result, yet it is exorbitant for k>1, and it is exceptionally hard to the information owner. To handle the requirements of VD-kNN, T-kNN, a strategy that works by preparing encrypted Delaunay triangulations is being presented that underpins the estimation of k. Also, this support diminishes the heap of the information proprietor. This gives exact results for k=1, yet when k>1 the outcomes vary. It returns square measure which is extremely rough. Nonetheless, it is being demonstrated that this precision is high. An instrument is being presented for upgrading scrambled Voronoi outlines and Delaunay triangulations that permits to deliver these precisions effectively. In partner dynamic way, with changing datasets then Geo-Tagging the closest neighbour inquiry preparing in Untrusted Cloud Environment.

II. Related Work

A. LOCATION CLOAKING

This techniques replace the exact location of a user with a cloaking region (CR), typically of rectangular shape. To establish result correctness, the CR must enclose the actual user location. Furthermore, CRs must satisfy certain necessity dictated by a privacy paradigm, which expresses the privacy requirements of the user (e.g., spatial k-anonymity (SKA) [1] requires each CR to contain at least k distinct users). Regardless of the method used to develop the CR, query processing at the LS side is performed with respect to a rectangular region, as opposed to an exact user location. In effect, the result returned by the LS is a super-set of the actual query result.

B. PRIVATE INFORMATION RETRIEVAL

(PIR) methods rely on a cryptographic protocol to achieve query privacy [1]. In a pre-processing phase, the LS constructs the POI database into a data structure relevant to the supported type of query, and maps it to an ordered array *D[1...n]*. At runtime, a query is altered from a context-based (i.e., spatial) query to a query-by-index (i.e., return the ith item), according to the data organization which is known by the users. When a user wishes to get D[i], s/he creates an encoded query object *q(i).* Using a mathematical transformation, the LS calculates privately (i.e., without studying the value of i) the result *r(D,q(i))* and sends it back to the user. PIR protocols ensure that it is computationally tough for the LS to recover the value i from *q(i),* but at the same time the user can easily re-construct *D[i]* from *r*.

C. ORDER PRESERVING ENCODING

Order-preserving symmetric encryption (OPE) is an encryption scheme (aka. cipher) whose encoding function preserves numerical ordering of the plaintexts. OPE has a long period of history in the form of one-part codes, which are series of plaintexts and the corresponding ciphertexts, both arranged in alphabetical or numerical order so only a single copy is needed for profitable encoding and decoding. One-part codes were used, for example, during World War I [2]. The reason for new interest in such schemes is that they allow profitable range queries on encoded data. That is, a remote untrusted directory server is able to index the (sensitive) data it receives, in encrypted form, in a data structure that permits efficient range queries (asking the server to return ciphertexts in the directory whose decodings fall within a given range, say [a, b]). By “efficient” it means in time logarithmic (or at least sub-linear) in the size of the directory, as performing linear work on each query is prohibitively slow in practice for large directories. In fact, OPE not only allows convenient range queries but allows indexing and query processing to be done exactly and as conveniently as for unencrypted data, since a query just consists of the encryptions of a and b and the server can locate the desired key in logarithmic-time via standard tree-based data structures. Indeed, subsequent to its publication, [2] has been referenced widely in the database community, and OPE has also been suggested for use in in-network aggregation on encoded data in sensor networks and as a tool for applying signal processing methods to multimedia content protection.

D. K - ANONYMITY IN RELATIONAL DATABASES

Anonymity was first discussed in relational directories, where advertised data (e.g., census, medical) should not be linked to specific persons. Recent work has focused on K-anonymity as defined in [3]: a relation satisfies K-anonymity if every tuple is identical to at least K-1 other tuples with respect to a set of quasi-identifier traits. Quasi-identifiers are fields (e.g., date of birth, gender, zip code) that can be linked to publicly available data to identify individuals. Records with identical quasi-identifiers form an anonymized group. Two techniques are used to transform a relation to a K-anonymized one: suppression, where some of the tuples are removed and generalized, which involves replacing specific values (e.g., phone number) with more general ones (e.g., only area code). Both methods lead to knowledge loss. It shows that anonymizing a high-dimensional relation leads to unacceptable damage of information due to the dimensionality curse.

E. RAndom SPACE PERTURBATION (RASP) TECHNIQUE

RASP is one type of multiplicative perturbation, with a novel combination of OPE, scope expansion, random noise injection, and random projection. Let’s consider the multifaceted data are numerical and in multifaceted vector space. The directory has k searchable dimensions and *n* records, which makes a $d \times n$ matrix X. The searchable elements can be used in queries and thus should be indexed. Let x represent a *d*-dimensional record, $x \in R^d$. Note that in the *d*-length vector space R^d, the range query conditions are represented as half-space functions and a range query is rendered to finding the point set in corresponding polyhedron area described by the half spaces [4].

The RASP perturbation utilizess three steps. Its security is based on the existence of random invertible real-value matrix generator and arbitrary real value generator. For each k-dimensional input vector *x*,

1) An order preserving encryption (OPE) method [1], E_{ope} with keys K_{ope}, is applied to each dimension of *x*: $E_{ope}(x,K_{ope}) \in R^d$ to change the elemental distributions to normal distributions with each dimension’s value order still preserved.
2) The vector is then enhanced to *d+2 dimensions* as $G(x) = ((E_{opt}(x))^T ,1,v)^T$, where the *(d + 1)* - th dimension is always a 1 and the (d + 2) - th dimension, *v*, is drawn from a arbitrary real number generator *RNG* that generates random values from a tailored normal distributions..
3) The (d + 2)-dimensional vector is finally transformed to
$F(x, K = \{A, K_{ope}, RG\}) = A((E_{ope}(x))^T, 1, v)^T,$ (1)
where A is a $(d+2)\times(d+2)$ arbitrarily generated invertible matrix with $a_{ij} \in R$ such that there are at least two non-zero values in each line of A and the last column of A is also non-zero.

K_{ope} and A are shared by all vectors in the directory, but v is randomly generated for each individual vector. Since the RASP-perturbed data records are only used for indexing and supporting query processing, there is no need to recover the perturbed data.

F. CASPER

A novel framework that turns traditional geo-positional servers and query processors to provide anonymous service to their customers. In Casper, mobile users can enjoy location-based services without the need to reveal their private location information. Upon registration with [5] Casper, mobile users specify their convenient level of privacy through a user-specified privacy profile. A user privacy profile includes two arguments *k* and A_{min}. *k* indicates that the mobile user wants to be *k*-anonymous, i.e., not differentiable among other k users while A_{min} indicates that the user wants to hide her location data within an area of at least A_{min}. Large values fork and A_{min} indicate more strict privacy requirements. Casper mainly consists of two parts, namely, the location anonymizer and the privacy-aware query processor. The location anonymizer is a verified third party that acts as a middle layer between mobile users and the location-based database server in order to:

(1) Retrieve the exact location information from mobile users along with a privacy profile of each user,
(2) Obscure the exact location information into veiled spatial areas based on each user privacy profile, and
(3) Send the veiled spatial areas to the location-based directory server. The privacy-aware query processor is embedded inside the location-based directory server to tune its functionality to accord with anonymous queries and cloaked spatial areas rather than the exact location information. There are three novel query types that are supported by Casper:

(1) Private queries over public information, e.g., *"Where is my nearest coffee shop"*, in which the individual who issues the query is a private entity while the data (i.e., *coffee shops*) are public,
(2) Public queries over private data, e.g., *"How many cars in a certain area"*, in which a public entity asks about personal private locations, and
(3) Private queries over private data, e.g., *"Where is my closest restaurant"* in which both the person who issues the query and the requested data are protected. With this classification in mind, conventional location-based query processors can support only public queries over public data. Due to the absence of the exact location information at the server, the anonymous query processor gives a candidate list of answers instead of a single exact answer. Hence it proves that the candidate list is inclusive, i.e., has the exact answer, and is low, i.e., a high quality answer is given to the users.

G. IND-CPA

Ciphertext identicalness is a property of many encoding schemes. Intuitively, if a cryptosystem possesses the property of indistinguishability, then an adversary will not be able to distinguish pairs of ciphertexts based on the message they encode. The property of identicalness under chosen plaintext attack is considered a basic requirement for most arguably secure public key cryptosystems, though some strategies also provide identicalness under chosen ciphertext attack and robust chosen ciphertext attack. Indistinguishability under chosen plaintext attack is equivalent to the feature of semantic immunity, and many cryptographic proofs use these interpretations interchangeably [6].

For a probabilistic asymmetric key encoding algorithm, indistinguishability under chosen plaintext attack [7] (IND-CPA) is defined by the following game between an attacker and a challenger. For schemes based on computational security, the attacker is formed by a probabilistic polynomial time and it must finish the game and yield a *guess* within a polynomial number of time steps. In this interpretation E(PK, *M*) represents the encoding of a message *M* under the key *PK*:

(1) The contender generates a key pair *PK*, *SK* based on some security value *k* (e.g., a key size in bits), and publishes *PK* to the attacker. The contender retains *SK*.

(2) The attacker may perform a polynomially bounded number of encryptions or other operations.

(3) Eventually, the attacker outputs two specific chosen plaintexts $M_{0,} M_1$ to the challenger.

(4) The contender selects a bit *b* € {0, 1} uniformly at random, and sends the *challenge* ciphertext *C* = E(PK, M_b) back to the attacker.

(5) The attacker is free to perform any number of additional computations or encryptions. Finally, it outputs a guess for the value of *b*.

A cryptosystem is indistinguishable under chosen plaintext attack if every probabilistic polynomial time adversary has only a negligible "advantage" over random guessing. An adversary is said to have a negligible "advantage" if it wins the above game with probability $1/2 + \in(k)$, where €(k) is a negligible function in the security parameter *k*, that is for every (nonzero) polynomial function *poly()* there exists k_0 such that $|\epsilon(k)| < \left|\frac{1}{poly(k)}\right|$ for all $k > k_0$.

Although the adversary knows $M_{0,}$ M_1 and PK, the probabilistic nature of E means that the encryption of M_b will be only one of many valid ciphertexts, and therefore encrypting $M_{0,}$ M_1 and comparing the resulting ciphertexts with the challenge ciphertext does not afford any non-negligible advantage to the adversary.

While the above definition is specific to an asymmetric key cryptosystem, it can be adapted to the symmetric case by replacing the public key encryption function with an "encryption oracle", which retains the secret encryption key and encrypts arbitrary plaintexts at the adversary's request.

III. SYSTEM MODEL

The system consist of three parties : owner, client, and outsourced service provider (SP) i.e cloud.
However, In this system the owner and client are not disjoint entities. In fact, the data owner in our case can be seen as a set of service users, each of whom sends his or her encrypted location data to the server. The client is also one of these users, who fulfills the role of querying the user. The SP is representing a location-centric service.
Figure 1 depicts how the user query will be processed under the cloud service as here,users are firing their query over the internet based application where the cloud server is trying to retrieve the address of user query from its cloud

storage to retrieve the required data from the file server with its complete authenticity i.e the user required data is completely secured and then sending back the data to the cloud server then to user.For protecting the file from the unauthorized person, one need to apply different types of privacy homomorphic algorithms.

Homomorphic encryption has been used to provide a strong privacy protection for the sensitive data. Homomorphic encryption allows addition and multiplication without the need for decryption to be directly performed on cipher texts and that too without loss of generality. The popular Parlier's homomorphic encryption is being used.

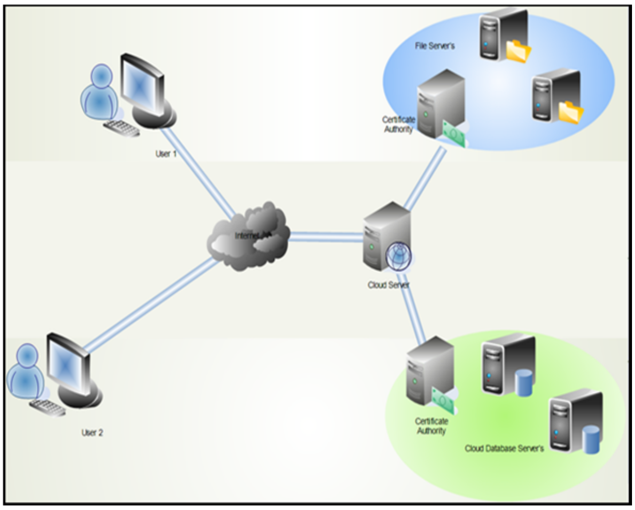


Figure 1 System Model

IV. System Modules

A. SPATIAL DATABASE MODULE

Spatial database is a database that is optimized to store and query data that represents objects de- fined in a geometric space. Most spatial databases allow representing simple geometric objects such as points, lines and polygons. Some spatial databases handle more complex structures such as 3D objects, topological coverages, linear networks, and TINs. While typical databases are designed to manage various numeric and character types of data, additional functionality needs to be added for databases to process spatial datatypes efficiently.

B. LOCATION PRIVACY MODULE

As mentioned previously, the dataset of points of interest represents an important asset for the data owner, and an important source of revenue. Therefore, the coordinates of the points should not be known to the server. So, it can be assumed that it is a honest-but-curious cloud service provider. In this model, the server executes correctly the given protocol for processing kNN queries, but will also try to infer the location of the data points. It is thus necessary to encrypt all information stored and processed at the server. To allow query evaluation, a special type of encryption that allows processing on cipher texts is necessary. Here, the mOPE technique is used from [8]. mOPE is a provably secure order-preserving encryption method, and techniques inherit the IND-OCPA security guarantee against the honest-but-curious server provided by mOPE. Furthermore, lets assume that there is no collusion between the clients and server, and the clients will not disclose to the server the encryption keys.

C. DATABASE OURSOURCING MODULE

The server receives the dataset of points of interest from the data owner in encrypted format, together with some additional encrypted data structures (e.g., Voronoi diagrams, Delaunay triangulations) needed for query processing. The server receives kNN requests from the clients, processes them and returns the results. Although the cloud provider typically possesses powerful computational resources, processing on encrypted data incurs a significant processing overhead, so performance considerations at the cloud server represent an important. The client has a query point Q and wishes to find the point's nearest neighbors. The client sends its encrypted location query to the server, and receives k nearest neighbors as a result. Note that, due to the fact that the data points are encrypted, the client also needs to perform a small part in the query processing itself, by assisting with certain steps.

D. VORONOI DIAGRAM

The first task is to find the 1NN of a query point. Voronoi diagrams [9] have been designed, which are data structures especially designed to support NN queries. An example of Voronoi diagram is shown in Figure 1. Denote the Euclidean distance between two points p and q by dist(p, q), and let P= p_1,p_2, p_n be a set of n distinct points in the plane. The Voronoi diagram (or tessellation) of P is defined as the subdivision of the plane into n convex polygonal regions (called cells)such that a point q lies in the cell corresponding to a point p(i) if and only if p(i) is the 1NN of q, i.e., for any other point p(j) it holds that dist (q,p(i)) < dist (q,p(j)) [1]. Answering a 1NN query boils down to checking which Voronoi cell contains the query point. In our system model, both the data points and the query must be encrypted. Therefore, the enclosure of a point within a Voronoi cell securely is checked. Next, such a secure enclosure evaluation scheme.

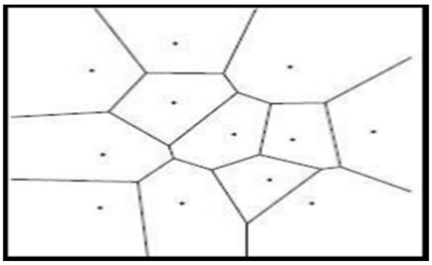


Figure 2 Voronoi diagram

Data Owner sends to Server the encoded Voronoi cell vertices coordinates, MBR boundaries for each cell, encoded right-hand side $R_{i,j}$ and encrypted $S_{i,j}$, for each cell edge. Client sends its encoded query point to the Server. Server performs the filter step, determines for each kept cell the edges that intersect the vertical line passing through the query point and sends the encrypted slope $S_{i,j}$, of the two edges to the Client. Client computes the left-hand side L $_{i,j}$, encodes it and sends it to the Server. Server finds the Voronoi cell enclosing the query point and returns result to Client.

E. k NEAREST NEIGHBOUR (kNN)

To support secure kNN queries, where k is fixed for all querying users, the VD-1NN method could be extended to generating order-k Voronoi diagrams. However, this method, VD-kNN, has several serious drawbacks:

(1) The complexity of generating order-k Voronoi diagrams is either O(k2nlogn) or O(k(nk)logn+ nlog3n), depending on the approach used. This is significantly higher than O(n log n) for order-1 Voronoi diagrams.

(2) That leads to high data encryption overhead than the data owner, as well as prohibitively high query processing time at the server (a k-fold increase compared to VD-1NN). Motivated by these limitations of VD-kNN, a secure distance comparison method (SDCM) has been introduced. Next, Basic kNN (B kNN) has been devised, a protocol that uses SDCM as building block, and answers kNN queries using repetitive comparisons among pairs of data points. B kNN is just an auxiliary scheme, very expensive in itself, but it represents the starting point for Triangulation kNN (T kNN), presented. T kNN builds on the B kNN concept and returns exact results for k=1. For k>1, it's an approximate method that provides high-precision kNN results with significantly lower costs.

V. IMPLEMENTATION

Java prototype which implements the data owner, the server and the client protocols has been developed. The Qhull library to generate order-1 Voronoi diagrams and Delaunay triangulations has been used. mOPE[8] was implemented using 32-bit encoding. The parallel computing section of the code was implemented using Java threads. The experimental testbed consists of an Intel i7 CPU machine with four cores.

The datasets of two-dimensional point coordinates ranging in cardinality from 200,000 to 1 million.

A uniform distribution of points in the unit space has been considered. In the case of processing on encrypted data, the actual data distribution has little or no effect on performance, since all values are treated in a similar way in encrypted form. Therefore, the results have been obtained for other distributions have been omitted. For encryption of slopes, used RC6.

The communication bandwidth for the wireless connection between the server and the client is set to 1Mbps.

The main performance metrics used to evaluate the proposed techniques are query response time and communication cost. The response time measures the duration from the time the query is issued until the results are received at the client. It includes the computation time at the server and the client, as well as the time required for transfer of final and intermediate results between client and server. Communication cost (measured in kilobytes) is important given that many wireless providers charge customers in proportion to the amount of data transferred.

In the setup phase, the data owner builds the Voronoi diagram or Delaunay triangulation for the dataset, encrypts these structures and sends them to the server. At runtime, there are two steps for each method: VD-kNN.

1) The client sends its encoded query point to the server which finds the Voronoi cells whose MBRs enclose the query point. For each of these cells, the server sends the encrypted slopes $S_{i,j}$ of two cell edges intersecting the vertical line passing through the query point.

2) The client computes the left-hand sides $L_{i,j}$

$$\mathbf{L}_{1,2} = \mathbf{y}_q\ \mathbf{S}_{1,2}\ \mathbf{*x}_q\ \mathbf{<1*S}_{1,2}\ \mathbf{*x}_1 + \mathbf{y}_1 = \mathbf{R}_{1,2}$$

and sends their ciphertexts to the server, which finds the Voronoi cell enclosing the query point. TkNN.

The values of x_q and y_q are variable for each query, but the Voronoi diagram does not change with the query, so, x_i, y_i, and $S_{i,j}$ remain constant. We denote the right-hand side and the left-hand side by Ri,j and Li,j, respectively. Ri,j is constant for a given query, and can be determined by the data owner when s/he uploads the database to the server. In addition, the data owner encrypts the value of slope Si,j with conventional encryption (e.g., RC6) and sends it to the server.

1) The client sends the encoded query square to the server, and the server finds the data points enclosed by the square. The server sends to the client the encrypted slopes $S_{i,j}$ of the perpendicular bisectors corresponding to each such data points. 2) The client computes the encoded left sides $L_{i,j}$

$$\mathbf{y}_q > \mathbf{S}_{i,j}\ \mathbf{*(x}_q\ \mathbf{x}_{i,j}\mathbf{)+y}_{i,j} \Leftrightarrow \mathbf{L}_{i,j} = \mathbf{y}_q\ \mathbf{S}_{i,j}\ \mathbf{*x}_q >> \mathbf{1*S}_{\mathbf{i},j}\ \mathbf{*x}_{i,j} + \mathbf{y}_{i,j} = \mathbf{R}_{i,j}$$

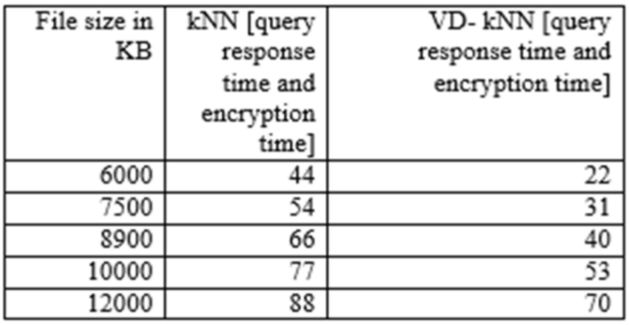
VI. RESULTS

Our framework first produces a system Voronoi outline and also border hubs of every cell as appeared. The pursuit can extend starting with one Voronoi cell then onto the next through the fringe hubs. We don't show associations or separation in the middle of values and their outskirt hubs because of the limited perception space. For straightforwardness, we utilize the quantity of catchphrases coordinated with given watchwords as the catchphrase pertinence in positioning score count. The numbers are shown alongside every POI and can offer clients some assistance with understanding how the two diameters, catchphrase significance and weighted separation, work in the outcome determination.

Four gatherings of execution assessment measurements are upheld; clients can change/explore:
- The extent of information datasets,
- The quantity of catchphrases,
- The quantity of hopefuls delivered by the arrangements,
and - The quantity of page access

The primary execution measurements used to assess the proposed systems are question reaction time and encryption time. Table.1 Show the reaction time measures the length of time from the time the question is issued until the outcomes are gotten at the customer. It gives the calculation time at the server and the customer, still on the grounds that the time required for exchange of last and transitional results in the middle of customer and server.

Table 1: kNN vs VD-kNN



File size in KB	kNN [query response time and encryption time]	VD- kNN [query response time and encryption time]
6000	44	22
7500	54	31
8900	66	40
10000	77	53
12000	88	70

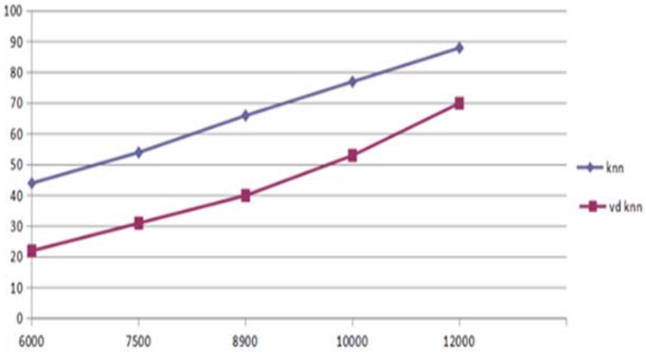


Figure 3: Response Graph

VII. CONCLUSION

A secure traversal framework in the indexed environment is given to secure protocols for such classic queries. The assumptions and approaches considered in this paper are thoroughly useful, dynamic to perform and effectively used under settings of different specification. It has been summarized that the process mentioned here, on privacy homomorphism, is used to protect processing queries on the cloud is highly scalable.

VIII. FUTUREWORK

All the schemes and the proposed schemes use different encryption strategies. However, they can be integrated on a single encrypted database. Range and kNN queries can be performed on balances and locations independently. The possibility of integrating different schemes in the VD-kNN model to support a wide range of applications makes EDBMS a practical solution to service out- sourcing. In the integrated solution, there is a need to ensure that the different schemes are aligned on the same security level.

REFERENCES

[1] Gabriel Ghinita, Panos Kalnis, Murat Kantarcioglu, and Elisa Bertino, “A Hybrid Technique for Private LocationBased Queries with Database Protection”, SSTD09

[2] Gabriel Ghinita, Panos Kalnis, Murat Kantarcioglu, and Elisa Bertino, “Approximate and exact hybrid algorithms for private nearest-neighbor queries with database protection”, Geoinformatica11

[3] Panos Kalnis, Gabriel Ghinita, Kyriakos Mouratidis, and Dimitris Papadias, “Preventing Location-Based Identity Inference in Anonymous Spatial Queries “

[4] Xu, Shumin Guo, Keke Chen, “Building Confidential and Efficient Query Services in the Cloud with RASP Data Perturbation” Huiqi Data Intensive Analysis and Computing Lab Ohio Center of Excellence in Knowledge Enabled Computing Department of Computer Science and Engineering Wright State University, Dayton, OH 45435

[5] Mohamed F. Mokbel Chi-Yin Chow Walid G. Aref, “The New Casper: Query Processing for Location Services without Compromising Privacy” Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN, Department of Computer Science, Purdue University, West Lafayette, IN

[6] Alexandra Boldyera, Nathan Chenette, Younholee and Adam O’Neil, “Order Preserving Symmetric Encryption”, GA USA

[7]https://en.wikipedia.org/wiki/Ciphertext_indistinguishability#Indistinguishability_under_chosen plaintext_attack_.28IND-CPA.29

[8] Gabriel Ghinita, Panos Kalnis, Ali Khoshgozaran, Cyrus Shahabi, and Kian-Lee Tan, “Private Queries in Location Based Services: Anonymizers are not Necessary”, SIGMOD08

[9] W. K. Wong, David W. Cheung, Ben Kao, and Nikos Mamoulis, “Secure kNN Computation on Encrypted Databases”, SIGMOD09