Efficient K-Nearest Neighbours Discovery using Complex Secure Evaluation Function in Geo Tagged Dataset

Kommuri Manoj
M.Tech-Computer Science & Engineering
Kallam Haranadha Reddy Institute of Technology,
Guntur, Andhra Pradesh, India

B. Tarakeswara Rao
Professor-Computer Science & Engineering
Kallam Haranadha Reddy Institute of Technology,
Guntur, Andhra Pradesh, India

ABSTRACT
The materialization of mobile equipment with fast Internet connectivity and geo-tagged capabilities has led to a revolution in personalized position-based services where clients are enabled to admittance information about points of interest that are pertinent to their interests and are also close to their geo tagged coordinates. The important type of queries that absorb location attributes is symbolized by nearest-neighbor queries, where a client wants to retrieve the k-Point of interests that are nearest to the user's current location. Entities specialized in various areas of interest and gather large amounts of geo-tagged data that appeal to subscribed users. In this paper the proposes development and implementation of frame work to provide more complex protected evaluation method on cipher texts, like skyline queries and security protection guarantees against the client, to prevent it from learning anything other than the received k query results.

Keywords
points of interest (POI), geo-tagged data, Query processing, and KNN Query.

1. INTRODUCTION
Now days the usage of portable devices with high speed network connectivity along with Geo positioning system greatly increased. With this equipment user can able to access information such as restaurants, colleges, hospitals, amusement parlors etc. Co-ordinate information of GPS is called point of interest (POI). The processes of extracting these POI's using queries are nearest neighbor query. In some cases user may know point of interest with limiting factor are called KNN queries. The following are the challenges first; the system is capable of protecting geo-tagged data from the usage of other purposes. Second it is capable of providing services to small organizational data owners offered as component plugging. Third it is capable of protecting sensitive geo-tagged data from anti social elements such as terrorists and Maoists etc. Fourth it is capable providing point's interests requested by the user is pay bill mode and should provide exactly K-nearest neighbors. In this paper the proposed development and implementation of frame work to provide more complex protected evaluation method on cipher texts, like skyline queries and security protection guarantees against the client, to prevent it from learning anything other than the received k query results.

2. RELATED WORK
A security detail is a vital requirement for cloud search approach. There are two methods for protecting location or point of interest in the specific area of user request for location tagged item. The main goal here is apply the requesting (query), the clients to extract their near point of interest with the confidentiality of user location the first related work is to implementation using Cloaking regions [8,9], but the majority of the solutions using the relating anonymity method and this not that much protective outliers[3,4,5]. The second approach related to proposed work is to use private information retrieval (P.I.R) [6,7]. The protocols of PIR are expensive in case of computational cost. In the work of Khoshgozaran etal [10] proposed an approach to find nearest neighbors using space transformation. But this approach suffers with problems accuracy declaim and decryption vulnerable. The work [2] of wong etal. Uses matrix transformation approach to encrypt the date points. If is vulnerable to the plain text attack.

3. PROBLEM DESCRIPTION
To process K-Nearest Neighbour query it requires three basic entities

1) The owner of the data set.
2) Client i.e. the User who request for KNN query results.
3) The cloud server as shown in following figure.

Let us assume that there are a different Geo tagged (2D) points of interest (POI) due to the lack of computational power for processing K-nearest neighbour (KNN) query , the owner of geo tagged data set out sources the task of storing and querying services to the cloud provides. The Geo tagged data resource is a valuable thing so that the system has to provide communication (query, results) in encrypted form among server, client, and owner.

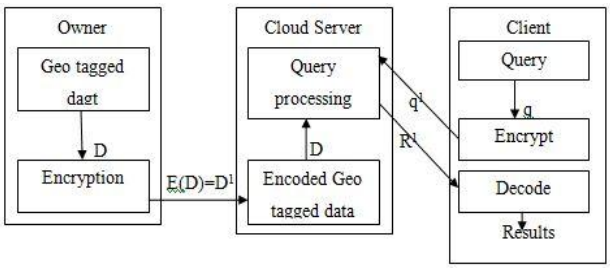

Figure.1

The transaction among three entities of system i.e. owner, server and client described as follows.

Step 1: The server receives Geo tagged point of interest in encrypted form (D1) from data owner.

Step 2: Now the server is ready to take the KNN query requests from clients.

Step 3: Client sends encrypted Query Q1 to server in order to get KNN query results

Step 4: Server receives the request (Q1) from client and result R1 apply directly to D1 and produce.

Step 5: Server sends encoded results (R1) to the client.

Step 6: Finally client decode the Results received from server.

To improve the query processing, it is implemented using complicated secure processing functions (skyline queries).

4. PROPOSED FRAMEWORK

This section mainly describes frame work and algorithm user for processing KNN query.

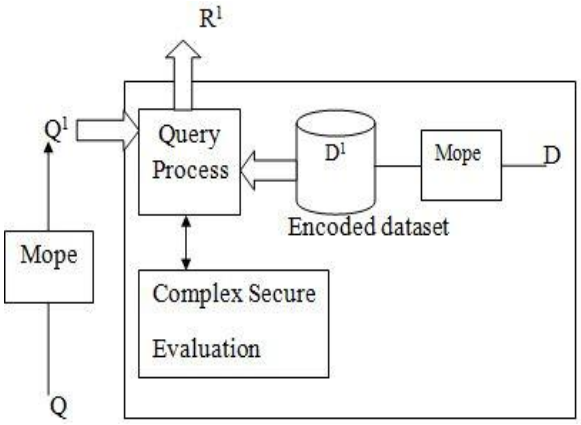

Figure.2

4.1 Algorithm

kNN Protocol skyline

1. Data Owner send s to Server: all encoded data points, and for each pair of points the encoded right-hand side RHm,n and encrypted slopes Mm,n.

2. Client sends the encoded query to Server.

3. Server finds the data points in the large query square and sends their AES – encrypted slopes Mm,n to client.

4. Client computes the encoded left-hand side LHm,n of Eq. (11) and sends them to Server.

5. Server returns k result points to Client.

4.2 Skyline Queries

Input: instance r with schema A = {A1, . . . ,Ad}

Output: skyA(r)

1: Partition r using the median mi of some attribute Ai. Let SHi = {t 2 r.ti (mi} and

SLi = r \ SHi ;

2: Compute skyA(SHi) and skyA(SLi) by recursively applying step 1;

3: Merge skyA(SHi) and skyA(SLi), i.e., determine TLi skyA(SLi) s.t. skyA(r) = skyA(SHi)[TLi] .

4.3 Performance

Table.1

method	Generation time of data	Encoding time	Query encryption time at client	Knn query processing time at cloud server
VD-kNN	K^2nlogn or k(n-k) + $n\log^3 n$	7kn	O(1)	kn
BkNN	n/a	n(n-1)/2	O(k(k-1) /2)	n(n-1)/2
TkNN	nlogn	5n	O(k)	n
Proposed method	nlogn-n	4n	o(k)	n

5. EXPERIMENTAL SETUP AND RESULTS

The proposed approach implemented using the integrated environment (IDE) as visual studio 2010 and coding Language C#. The SQL Server 2008 can be used as server data base using the machine configuration of intel dual core with 2 GB Random access memory and Windows 7 operating system minimum requirements of band width for connection among entities of the proposed system is 1 Mbps. The implemented simulated environment to test the approach 600 X 600 points in dot net form. Simulated environment tested with Range of different parameters.

Table.2

Parameter	Range
Width	200-600
Height	200-600
K	5-15
Key size	16-32 bits
# point of interest	2K-5K

The following graphs depict comparative performance with existing techniques.

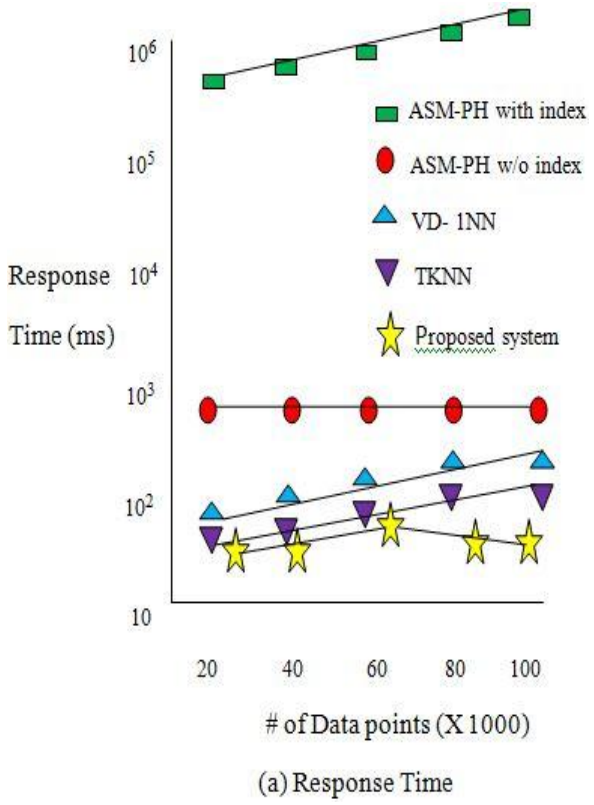

Figure.4

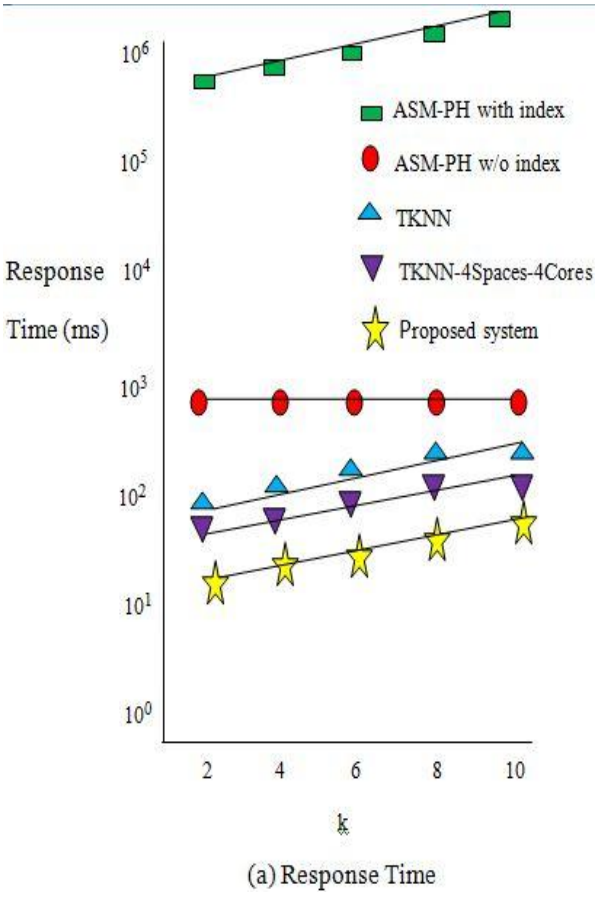

Figure.5

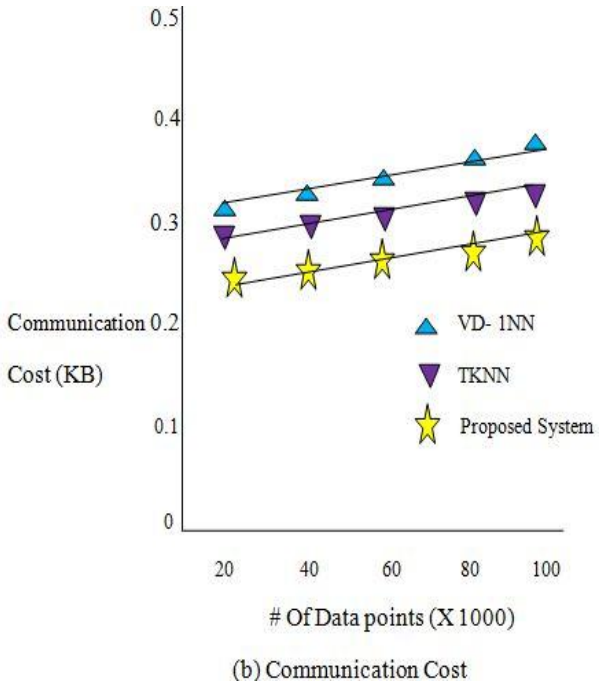

Figure.6

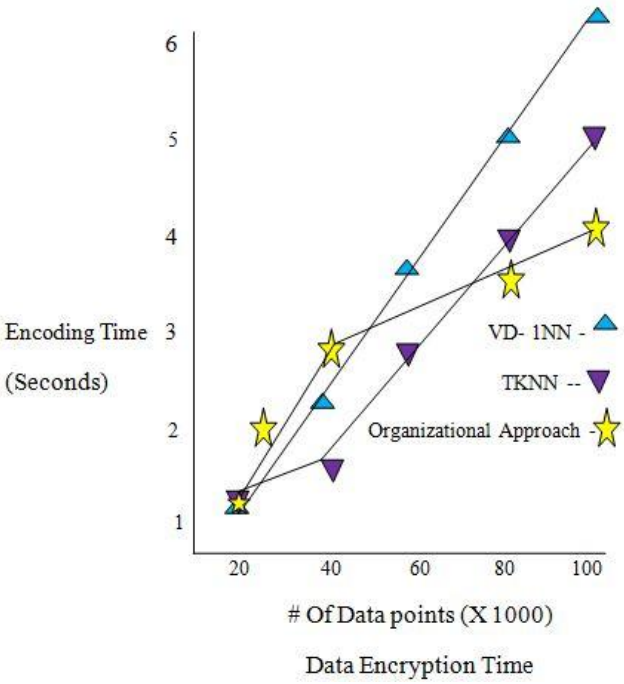

Figure.7

6. CONCLUSION

The proposed scheme to support secure k nearest neighbor query process using mutable order-preserving encoding as building block and development and implementation of frame work to provide more complex protected evaluation method on cipher texts, like skyline queries and security protection guarantees against the client, to prevent it from learning anything other than the received k query results.

7. REFERENCES

[1] sunoh choi, gabriel ghinita, hyo-sang lim, and elisa bertino, "secure knn query processing in untrusted

cloud environments" transactions on knowledge and data engineering, vol. 26, no. 11, november 2014

[2] W.K. Wong, D.W. Cheung, B. Kao, and N. Mamoulis, "Secure kNN Computation on Encrypted Databases," Proc. ACM SIGMOD Int'l Conf. Management of Data, 2009.

[3] H. Hu, J. Xu, C. Ren, and B. Choi, "Processing Private Queries over Untrusted Data Cloud through Privacy Homomorphism," Proc. IEEE 27th Int'l Conf. Data Eng. (ICDE), 2011.

[4] H. Xu, S. Guo, and K. Chen, "Building Confidential and Efficient Query Services in the Cloud with RASP Data Perturbation," IEEE Trans. Knowledge and Data Eng., vol. 26, no. 2, pp. 322-335, Feb. 2014.

[5] B. Yao, F. Li, and X. Xiao, "Secure Nearest Neighbor Revisited," Proc. IEEE 27th Int'l Conf. Data Eng. (ICDE), 2014.

[6] G. Ghinita, P. Kalnis, A. Khoshgozaran, C. Shahabi, and K.-L. Tan, "Private Queries in Location Based Services: Anonymizers Are Not Necessary," Proc. ACM SIGMOD Int'l Conf. Management, 2008.

[7] G. Ghinita, P. Kalnis, M. Kantarcioglu, and E. Bertino, "Approximate and Exact Hybrid Algorithms for Private Nearest- Neighbor Queries with Database Protection," J. Geoinformatica, vol. 15, pp. 699-726, 2011.

[8] M. Gruteser and D. Grunwald, "Anonymous Usage of Location- Based Services Through through Spatial and Temporal Cloaking," Proc. First Int'l Conf. Mobile Systems, Applications and Services, 2003.

[9] P. Kalnis, G. Ghinita, K. Mouratidis, and D. Papadias, "Preserving Location-Based Identity Inference in Anonymous Spatial Queries," IEEE Trans. Knowledge and Data Eng., vol. 19, no. 12, pp. 1719- 1733, Dec. 2007.

[10] A. Khoshgozaran and C. Shahabi, "Blind Evaluation of Nearest Neighbor Queries Using Space Transformation to Preserve Location Privacy," Proc. 10th Int'l Conf. Advances in Spatial and Temporal Databases, 2007.