

# Higher Confidentiality through Grouping Hilbert and Voronoi over Validation of K-Nearest Neighbor Query on Spatial Network

V. Srilakshmi<sup>1</sup>

Anna University, Computer Science and Engineering  
Akshaya College of Engineering and Technology  
Coimbatore, India

P. Dhamodharan<sup>2</sup>

Anna University, Computer Science and Engineering  
Akshaya College of Engineering and Technology  
Coimbatore, India

**Abstract—** The administration of transshipment systems has become increasingly important in many applications such as position-based services, supply cycle management, travel control, and so on. These applications usually involve queries over spatial networks with vigorously changing and problematical travel conditions. There may be possibilities of user's privacy violated when they are querying about the location information on the third party servers where the location information about the users will be tracked. The malicious attackers may steal the location information about the users. The k nearest neighbor query verification with location points on Voronoi diagram increases the verification cost on mobile clients. The reverse nearest neighbor queries by assigning each object and query with a safe region is applied such that the expensive recomputation is not required as long as the query and objects remain in their respective safe regions. The proposed system reduces the communication cost in client-server architectures because an object does not report its location to the server unless it leaves its safe region or the server sends a location update request. Hilbert curve is used here for the capability of partially retaining the neighboring adjacency of the original data. The user data is protected by applying Hilbert transform over the original values and storing the transformed values in the Hilbert curve. The experimental results prove that the proposed work, Hilbert transform based reverse nearest neighbor provides the better result in terms of time complexity, memory consumption and VO size than the existing work.

**Keywords—** Hilbert Curve, Voronoi diagram, Hilbert Transform.

## I. INTRODUCTION

Now, in the fast growing world the users are highly dependent on the location based services. The location information is extracted from spatial databases stored in the third party server using a spatial query. The antisocial

elements of world with the help of the technologies hack the location information of the user from the server there by no security support for the user data. The data security lacking leads to critical problem such as misusing the personal information of an individual. In these circumstances in order to provide security to users, existing system uses Network Voronoi diagram. The Network Voronoi provides security at minimum level. To provide increased security the proposed work applies Hilbert Transform algorithm over the Network Voronoi as a result high level of security is provided to user data.

Searching for concealed patterns from large data storage locations is formalized as Data mining. SDM(Spatial Data mining) is the process of discovering fascinating and unexplored, potentially useful patterns from large spatial data warehouses. SDM is used to find implicit and explicit regularities, relations between spatial data and non-spatial data.

Spatial data is a data related to the location and spatial dimensions of geographical items. Spatial data is also known as geospatial data or geographic information. Spatial data exists in the form of raw fact, complete information that determines the geographic location of features and boundaries such as natural or constructed features ocean and more. Spatial data is usually stored as coordinates and topology and that data can be mapped. Spatial data describes information related to the space occupied by objects. The data consists of geometric information and can be either discrete or continuous. Discrete data might be a single point in multidimensional space and Continuous data spans a region of space.

Spatial databases are database systems that stores and handles spatial data. They are designed to handle both spatial attributes and the non-spatial attributes of that data. A Spatial database is a database that is optimized to store and query data related to objects in space, including points and polygons. The very large size of spatial databases also requires additional techniques for manipulating and cleaning the data in order to prepare it for analysis.

A versatile ecosystem is created for radical change of the way geospatial data are stored, managed, served, and shared. The ecosystem is formed by combination of Cloud-based Solutions and Mobile devices and is also called as database outsourcing; Outsourcing is termed as paying to some authority for finishing a particular work. The authority in turns appoints people for the completion of the work. The Data Possessor (DP) delegates the management of its database to a third-party Cloud Access Provider, the Access Provider (AP) is responsible for indexing the data, answering client queries, and updating the data on requests from the DP's. Mobile consumers, which are used to send their queries to DP's, now submit queries to AP and retrieve results from AP directly. For example, Microsoft Bing Maps partners with NAVTEQ, a major provider of base electronic navigable maps, to provide web mapping services for the public. In this case, NAVTEQ is the DP, and Bing Maps is the access provider (AP). Cloud computing provides flexible resources that can easily be scaled up or down based on user demands, effectively reducing the operational and maintenance expenses for Data Possessor[1].

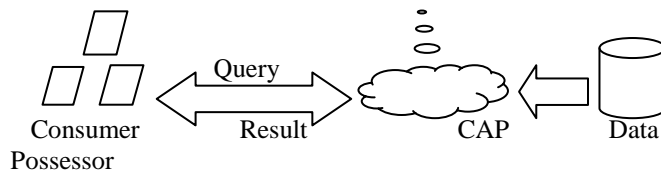


Figure 1 Database Outsourcing Architecture

Most interesting topic in field of research is cloud computing. Due to its wide variety of features and a special feature named as “any time any where access” it sounds high in the field of Research and Development. Although database outsourcing provides data possessor with a more efficient, economical, and flexible solution, it also introduces new concerns. The query veracity concern means to ensure that the query results returned by AP are still reliable. As Access Provider AP is not the real possessor of the data, it might return incorrect results to mobile consumers out of its own interests, for example, an AP which hosts collection of cafés might favour some that pay more commercial fees. Moreover, an AP might return suboptimal results to query clients by applying flawed or substandard algorithms in order to save estimation resources. Therefore, providing a mechanism that allows clients to authenticate the precision and completeness of the query result is necessary. Specifically, correctness means all data returned by SP originate from DP without any forgery and the query result is matching to that computed by DP. Completeness means all qualified results have been included by SP in the result set.

## II. LITERATURE SURVEY

### A. Authentication of K Nearest Neighbor Query on Road Networks

Outsourcing spatial databases [1] to the cloud provides an economical and flexible way for data owners to deliver spatial data to users of location-based services. However, in the database outsourcing paradigm, the third-party service provider is not always trustworthy, therefore, ensuring spatial query integrity is critical. Therefore, providing a mechanism that allows clients to authenticate the correctness and completeness of the query result is necessary.

The existing approaches cannot verify both the distance and the shortest path to the K-NN results simultaneously, the proposed uses a network Voronoi diagram-based verification approach that utilizes the network Voronoi cell of each result object to verify the correctness and completeness of the K-NN result with regard to both distance and path.

### B. Spatial Query Integrity with Voronoi Neighbors

With the popularity of location-based services [12] and the abundant usage of smart phones and GPS-enabled devices, the necessity of outsourcing spatial data has grown rapidly over the past few years. Meanwhile, the fast arising trend of Cloud storage and Cloud computing services has provided a flexible and cost-effective platform for hosting data from businesses and individuals, further enabling many location-based applications.

Nevertheless, in this database outsourcing paradigm, the authentication of the query results at the client remains a challenging problem. The existing method focus on the OSDB model and propose an efficient scheme, called VN-Auth, which allows a client to verify the correctness and completeness of the result set .

### C. Indexing Network Voronoi Diagrams

The Network Voronoi diagram [2] and its variants have been extensively used in the context of numerous applications in road networks, particularly to efficiently evaluate various spatial proximity queries such as K-NN and closest pair. Although the existing approaches successfully utilize the network Voronoi diagram as a way to partition the space for the specific problems, there is little emphasis on how to efficiently find and access the network Voronoi cell containing a particular point or edge of the network.

On using index structures on network Voronoi diagrams that enables exact and fast response to a query in road networks. The existing index structures, treats a network Voronoi cell as a simple polygon, and yield inaccurate results due to the network topology, and fail to scale to large networks with numerous Voronoi generators. Voronoi-Quad-tree is used to overcome the drawbacks of the existing method.

#### D. Authenticated Multistep Nearest Neighbor Search

The importance of authenticated query [6] processing increases with the amount of information available at sources that are untrustworthy, unreliable, or simply unfamiliar. This is the first work addressing authenticated retrieval from such sources using the multi-step k-NN framework. The direct integration of optimal NN search with an authenticated data structure incurs excessive communication overhead. To overcome communication overhead C-AMN, a technique that addresses the communication-specific aspects of NN can be used and it also reduces the overhead that occurs due to transmission overhead.

#### E. Partially Materialized Digest Scheme

In the outsourced database model, [10] a data owner publishes the database through a third-party server; the server hosts the data and answers user queries on behalf of the owner. Since the server may not be trusted, or may be compromised, users need a means to verify that answers received are both authentic and complete, that the returned data have not been tampered with, and that no qualifying results have been omitted.

A result verification approach for one dimensional query, called PMD that applies to both static and dynamic databases can be used. PMD uses separate indexes for the data and for their associated verification information, and only partially materializes the query.

#### F. Query Access Assurance in Outsourced Databases

Query execution assurance [15] is an important concept in defeating lazy servers in the database as a service model. The process of extending query execution assurance to outsourced databases with multiple data owners is highly inefficient. To cope with lazy servers in the distributed setting, the proposed QAA that focuses on I/O-bound queries is used.

The goal in QAA is to enable clients to verify that the server has honestly accessed all records that are necessary to compute the correct query answer, thus eliminating the incentives for the server to be lazy.

### III. OUTLINE OF THE PROJECT

The key idea of the project is to advance the security of the user and to focus on efficient retrieval of solutions. The paper focus on a special kind of data mining called as spatial data mining. Generally when the user surfing for certain information over the online network. The location information of the user will be stored in the third-party server.

As many of the third party server not secured there is possibility of hacking the information by the hacker. Thus information about location of the user gets leaked that leads to several critical problems such as kidnapping the particular individual. In order to provide high level security to user .The information about users can be stored in the Network Voronoi diagram over which Hilbert

Transformation is applied to increase security to the user information.

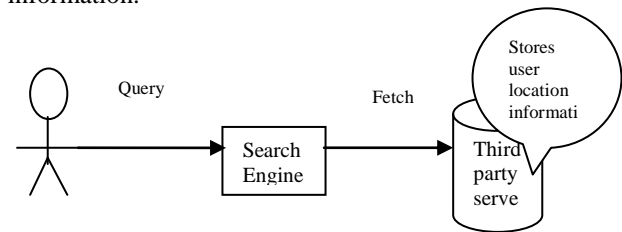


Figure 2 Storing User Information

In the case of Query Retrieval Process the results retrieved should be clear and complete and should not contain duplicates. The existing system uses K-NN for query retrieval process which includes redundant data.

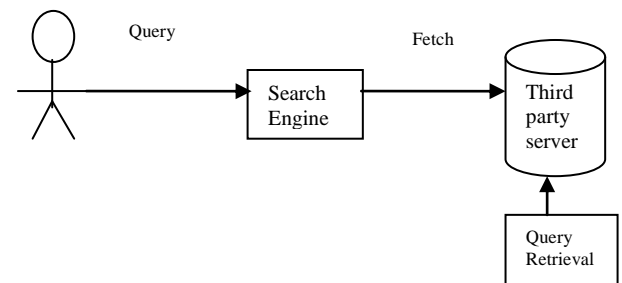


Figure 3 Query Retrieval

### IV. PROPOSED MODEL

The Voronoi diagram is used in the existing work to represent the location information in the graphical format. The K-NN classification algorithm is utilized on the Voronoi diagram for retrieving the location data as per user demand. The user data may get leaked, because of less security in the Voronoi diagram and also the K-NN classification algorithm cannot provide the accurate nearest location information to the users. To overcome these troubles the reverse RK-NN and Hilbert transformation are implemented to provide a solution to the problems.

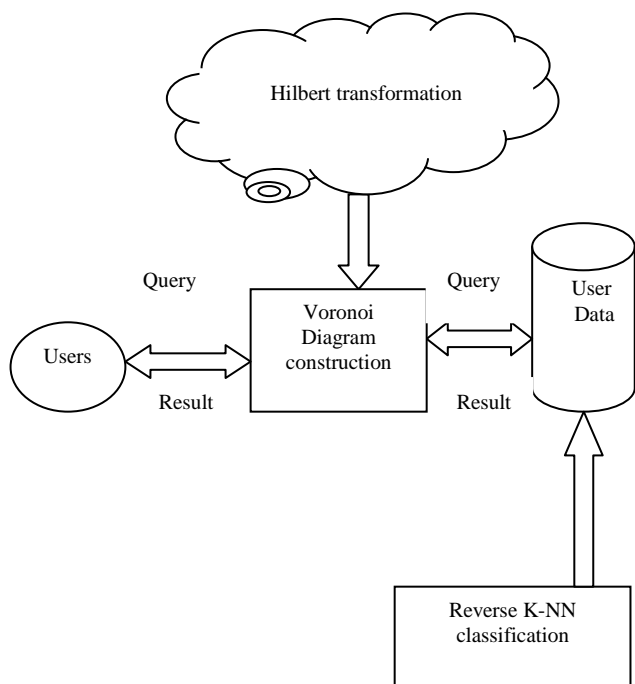


Figure 4 Architecture Diagram

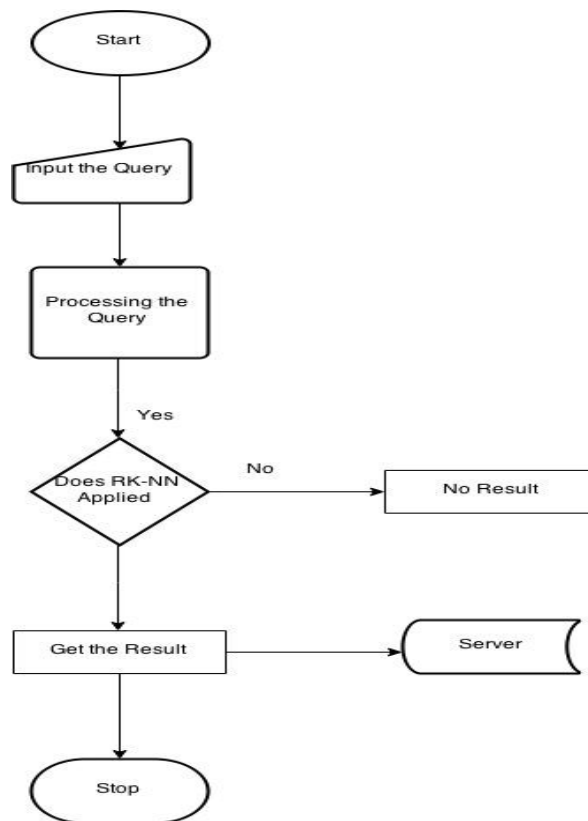


Figure 5 Flow diagram for User Query

**Step1: Key Generation / Query Submission**

Initially, the keys are generated and distributed over an network. DO obtains a private and a public key through a key distribution centre. The private key is confidential and is accessed only by DO, whereas the public key is accessible by all clients. Using its private key, the DO sends the data to SP which is used for query processing. The queries are gathered from the user and based on query the data processing take place on spatial networks and the queries will be then authenticated by using the RK-NN classification algorithm by getting the details from the neighbors.

**Step2: Forming the Spatial Network**

The creation of users, servers, and the query processor forms the network. The network formation is done by using the java platform. The user is the one who queries to the network for gathering the location information. The server is the one who stores and gives the location based information present in their storage. The server stores the spatial network information in the format of Voronoi diagram. The query processor is the one who gathers the queries from the user and will retrieve the results for those queries from the servers.

The below flow chart describes how results are obtained for the user query.

The previous technique used for storing the information is the MBR. The MBR represents the objects in form of circular points within the rectangular boundary. The objects are spatially distributed and exact location of the object within the boundary is easily identified, leads to security lacking for the object. The Voronoi diagram on comparing with MBR provides high security by storing the data within an irregular polyhedron structure. The exact location of the object cannot be determined due to unstructuredness of the Voronoi. Consider a set of distinct objects say  $P = \{p_1, p_2, \dots, p_n\}$  in a region  $R$ , the Voronoi diagram of  $P$ , denoted as  $VD(P)$ , partitions the space of  $R$  into  $t$  disjoint regions, such that each object  $p_i$  in  $P$  belongs to only one region and every point in that region is closer to  $p_i$  than to the other objects of  $P$ . The region around  $p_i$  is called the Voronoi polygon or Voronoi cell of  $p_i$ , denoted as  $VC(p_i)$ . Therefore, the Voronoi diagram of  $P$  is union of all Voronoi cells  $VD(P) = \{VC(P_1), VC(P_2), \dots, VC(P_t)\}$ . Voronoi neighbors shares a common edge.

**Step3 Applying Hilbert Transformation over Network Voronoi Diagram**

Hilbert curve is a space filling curve that is used only to find shortest path to reach the destination. The continuous research process leads to a solution that by applying the Hilbert curve along with transformation over the Network Voronoi diagram provides high confidentiality to the user. The user data remains protected.

Hilbert curve is a specialised curve that is highly complex in this structure that leads to complex index calculation. Hilbert transformation is used to store range of values in the curve over the Voronoi diagram. The transformation prevents the unauthorized user from getting the exact value.

Thus the security of the user is ensured and thereby preventing the leakage of the highly protected information and in today modern environment security plays important role in different fields. Security must be ensured in every day today activities of the user.

#### Step 4 Query Retrieval Process

The process deals with retrieval of results for an input query. The paper focus on spatial mining and it deals with getting spatial data for a spatial query. The existing system uses K-NN (K-Nearest Neighbor) to get K-NN spatial data results for the inputted spatial query.

K-NN has few deficiencies in processing query such as 1) Highly Dependent on Training data 2) Includes Redundant data 3) Increased Processing time 4) Low speed. The above drawback leads to inefficient query processing.

The proposed system uses RK-NN (Reverse K-Nearest Neighbor) technique for retrieving spatial results for the given spatial query. The bidirectional K-NN (RK-NN) take search to the next level and produces accurate results. The proposed algorithm over comes the drawback of the K-NN.

The advantages of RK-NN includes 1) Eliminates redundant data 2) Less processing time 3) High speed 4) Less memory consumption. The proposed system provides better results when compared to the existing system in terms of high security and efficient query retrieval process.

The below flow chart describes about the security of the user the data gets protected when stored over the Hilbert transform with Network Voronoi on the Server.

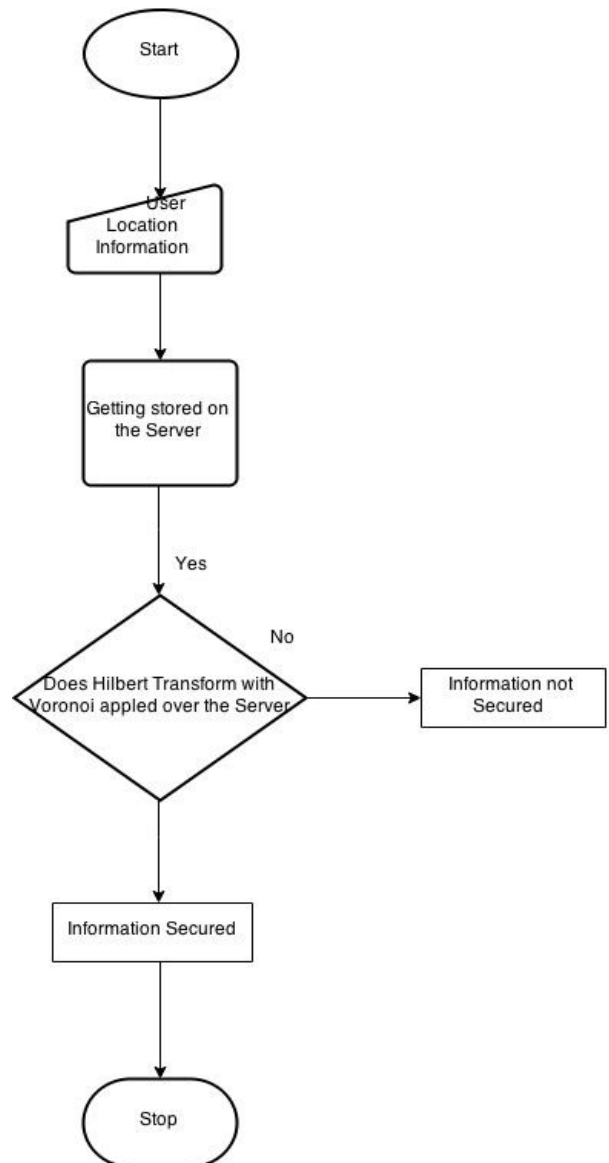


Figure 6 Flow diagram for Authentication

## V. PERFORMANCE EVALUATION

The performance evaluation is an important task that focuses on evaluating a system ability based on the performance metrics. The proposed system on comparing with existing system provides better results based on the parameters such as time, memory and size.

The analyzation and comparison of the performance offered by the existing system and the proposed system is an essential task. The performance is evaluated by the parameters such as terms time complexity, memory complexity and Voronoi size. Based on the comparison and the results from the experiment shows that proposed approach works better than the existing system.

The verification object is the one which is used to authenticate the query sent by the users. The size of the VO should be less which will increase the performance level. Here on comparing the existing system and proposed system in terms of VO size.

Proposed system has small VO size compared to the existing system. The VO means the Voronoi and it is the region where the all the information get stored on the server. The compact size of the Voronoi increases the security and reduces the intrusion of the data by the hackers (unauthorised people).



Figure 7 VO Size Comparison

In the above graph, comparing the VO size of the proposed system with the existing technique. In the graph, x axis will be the techniques and y axis will be VO Size in KB bytes. From the graph it is easily understand that the proposed system has small VO size. So the proposed system has smaller/better VO size which is compared to the existing system.

The proposed system is evaluated in terms of the time complexity in other words computation time of the query responding with the existing system and proposed system. It is defined as the time taking for the query response process.

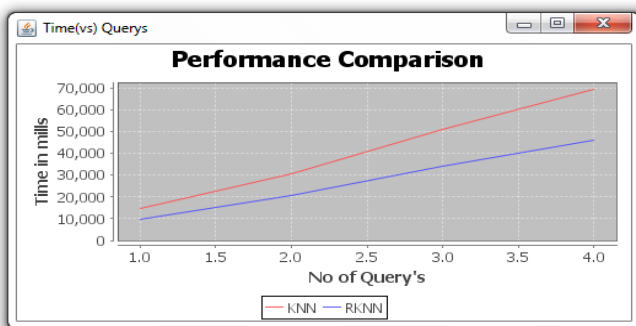


Figure 8 Time Complexity

In the above graph, comparing the time complexity of the proposed system with the existing technique. In the graph, x axis will be the methods K-NN and RK-NN and y axis will be time in milli seconds. From the graph it is easily understand that the proposed system consumes less time, so the proposed system has low time complexity compared to the existing system.

Memory consumed for processing each and every query by both existing K-NN and the proposed RK-NN is compared in the graph.

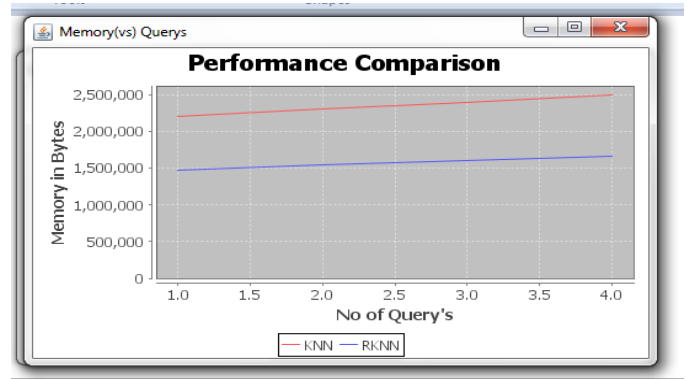


Figure 9 Memory Consumption

In the above graph, comparing the memory consumption of the proposed system with the existing technique. In the graph, x axis will be the methods K-NN and RK-NN and y axis will be memory in bytes. From the graph it is easily understand that the proposed system consumes less memory, so the proposed system has less memory consumption compared to the existing system.

## VI. CONCLUSION AND FUTURE SCOPE

In existing work, the query verification problem for k-nearest-neighbor queries on road networks in cloud environment is implemented. While existing approaches proposed in this domain cannot verify both the distance and the shortest path to the K-NN results simultaneously, a network Voronoi diagram-based verification approach that utilizes the network Voronoi cell of each result object to verify the correctness and completeness of the K-NN result is implemented with regard to both distance and path. To retrieve the better results than the K-NN classification algorithm, the RK-NN classification algorithm is implemented in the work. The user privacy is also guaranteed by introducing the method called Hilbert transformation in which the user's exact location information will be hidden from the hackers by transforming it to some other format. The performance evaluation conducted proves that the proposed method assures better result in terms of time complexity, memory consumption and VO size than the existing work

In the future focused on implementing the experiment as an android application. By implementing as an android application it will be more useful for users navigating around the world.

## ACKNOWLEDGEMENT

The success of a work depends on the team and its cooperation. I take this opportunity to express my gratitude and sincere thanks to everyone who helped me in my project. First and foremost, I would like to thank the Management for the Excellent Infrastructure, facilities and the constant support provided for the successful completion of the project work.

I thank Associate Professor **MrP.Dhamodharan** M.E., (Ph.D), Head of the Department, Computer Science and Engineering, for his valuable guidance, continuous support and suggestions to improve the quality of the project work. I express my special thanks to my guide, Associate Professor, **MrP.Dhamodharan** M.E., (Ph.D) for his valuable guidance, insightful comments and continuous support to carry out the project work.

I express my deep sense of gratitude to all the Faculty members and supporting staff for their continuous support in completing this project work successfully. My sincere thanks to my family members and friends for their support and motivation to carry out this project successfully.

## REFERENCES

- [1] Cyrus Shahabi, Ling Hu, Wei-Shinn Ku, YananJing, "Authentication of k Nearest Neighbor Query on Road Networks," IEEE VOL. 26, NO. 6, JUNE 2014
- [2] U. Demiryurek and C. Shahabi, "Indexing network Voronoi diagrams," in Proc. 17<sup>th</sup> Int. Conf. DASFAA, Bussan, South Korea, 2012, pp. 526-537.
- [3] M..Erwing, "The graph Voronoi diagram with Applications," Network vol. 36, no. 3, pp. 156-163, Oct.2012.
- [4] Franceso Bonchi, George Kollios, Aristides Gionis, Michalis Potamias, "K-Nearest Neighbors in Uncertain Graphs," J.Comput.Secur., vol. 16 2012.
- [5] Haixun Wang, LinLiu, Bolin Ding, Ruoming Jin, "Distance-Constraint Reachability Computation in Uncertain Graphs," vol. 17, no. 1, pp. 50-92, 2011
- [6] P. Karras, D. Pappadias, S. Papadopoulos, L. Wang, Y. Yang, "Authenticated Multi step Nearest Neighbor Search," IEEE Trans. Knowl. Data Eng., vol 23, no. 5, pp. 641-654, May 2011.
- [7] G. Kollios, S. Papadopoulos, D. Pappadias, Y. Yang, "Spatial outsourcing for location based services," in Proc. IEEE 24<sup>th</sup> ICDE, Cancun, Mexico 2009, pp. 1082 - 1091.
- [8] K.C.K. Lee, W.C. Lee, B. Zheng, Y. Tian, "ROAD: A new spatial object search framework for road network," IEEE Trans. Knowl, DataEng., vol. 24, no.3, pp.547-560, Mar 2012.
- [9] C.E. Leiserson, T. H. Cormen, R..L. Rivest and C.Stein, "Introduction to algorithms", Cambridge, MA, USA: MIT PRESS, 2009.
- [10] K. Mouratidis, D. Sacharidis and H. Pang, "Partially Materialised Digest Scheme: An efficient verification method for outsourced databases," VLDBJ., vol. 18, no. 1, pp.363-381, 2009
- [11] H. Pang and K. L. Don, "Authenticating query results in Edge computing," in Proc ICDE Washington 2010.
- [12] C.Shahabi, S.Bakiras, L.Hu, W.S. Ku, "Spatial Query Integrity with Voronoi Neighbors," IEEE Trans. Knowl. Data Eng., vol. 25, no. 4, pp. 863-876, Apr.2013.
- [13] R..Sion, "Query execution assurance for outsourced databases," in Proc.31<sup>st</sup>INT. Conf. VLDB, Trondheim, Norway, 2009, pp.601-612
- [14] K.L. Tan and W. Cheng, "Query assurance verification for outsourced multidimensional databases," J. Comput. Secur., vol. 17, no. 1, pp. 101- 126, 2009.
- [15] Wango Le and Feifei Li "Query Access Assurance in Outsourced Databases", 2009.



## Authors Profile

V.Srilakshmi received the B.E. degree in Computer Science and Engineering from the P.A College of Engineering and Technology, Coimbatore, Anna University Chennai, India. Currently doing M.E. in Computer Science and Engineering in Anna University Chennai, India Her research interest includes Datamining, and Cloud Computing.



## Authors Profile

MrP.Dhamodharan received the M.E. degree in Computer Science and Engineering from Coimbatore Institute of Engineering and Technology and has 10 years of experience in Teaching. Currently doing Ph.D in Computer Science and Engineering in Anna University Chennai, India His research interest includes Datamining, Networking and Cloud Computing, Image Processing.