ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Vol. 4, Issue 1, January 2016

Review On-Processing K-NN Queries over Moving Objects through Distributed Strip Index and DKNN

Suvarna D Vahadne[1], Ambika D Metgar[2], Sonabai V Reddy[3], Ashwini V Talekar[4], Archana Chaugule[5]

Student, Dept. of CS, Pad.Dr.D Y Patil College of Engineering and Technology, Pimpri, Savitribai Phule Pune University, Pune, India[1,2,3,4]

Professor, Dept. of CS, Pad.Dr.D Y Patil College of Engineering and Technology, Pimpri, Savitribai Phule Pune University, Pune, India[5]

ABSTRACT: Most of the applications involving moving objects is the task of processing k-nearest neighbor (k-NN) queries. many existing methods to this problem are designed for the centralized setting where query processing do on a single server; it is difficult, for them to scale to a distributed setting to handle the large volume of data and concurrent queries that are increasingly common in those applications. To identify this problem, purpose a suite of solutions that can support scalable distributed processing of k-NN queries. first present a new index structure called Dynamic Strip Index (DSI), which can better to different data distributions than existing grid indexes .it can be naturally distributed across the cluster, therefore lending itself well to distribute processing to represent a distributed k-NN search (DKNN) algorithm based on DSI. The advances of GPS technology and wide-ranging usage of wireless communication devices provide the collection of large amount of spatiotemporal data of special attract are data related to moving objects. By modeling and analyzing moving objects data learn about the moving objects behavior and even predict their future locations. Discovery of patterns and assumption of future movement can greatly impact different fields. Devising the correct inference is a scientific and computational challenge. basics of moving objects' representation and Space partitioning are presented.

KEYWORDS: k Nearest Neighbor, Multiple KNN queries, spatio-temporal database, information sharing strategy, queries processing.

I. INTRODUCTION

All moving things in the real world can contain Spatio temporal data, containing time and space attributes simultaneously.Data that are moving in time with interchanges of their location or shape describe a moving object. In many applications, to know moving objects current location in advance is very appreciable. Discovery of patterns of future movement can greatly effect on different fields. Some of examples are analyzing wild animals' movement in order to predict their seasonalmovement from one region to another, monitoring of vehicles and analyzing their movement in order to control the traffic ,analyzing the movement of a mobile user roaming around and changing access points that assure level of quality of service in wireless network or analyzing and predicting the movement of aircraft in order to develop defending technique. There are also situations in which determining the accurate location of a moving object is not possible, for example when the moving object goes in a shadow area of GPS and estimate the location by referencing the previous location which were provided in visible regions, becomes mandatory. As technology advances, we focus more available data on moving objects, thus increasing our potential to mine spatiotemporal data. We can use these data to analyze moving objects behavior and to predict their future locations. k nearest neighbor (k-NN) queries over moving objects using dynamic strip index is a fundamental operation in many location-based applications.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Vol. 4, Issue 1, January 2016

II. RELATED WORK

The movement is continuous; GPS and communication technology allow us to sample an object's position that is to obtain the position at discrete instance of time. Obtaining these samples are extract object's movement. The simple approach is to use linear interpolation. The sampled positions turn out to be the end points of line segments of polyline and the motion of an object is represented by entire polyline in three-dimensional space. The movement of an object (i.e. the trace of an object) is called trajectory. Trajectories can have different characteristics depending on the characteristics of moving objects they represent, and depending on the application requirements. Two important general characteristics of trajectories are direction and speed of the movement. Trajectories correlate to pertaining spatial environment (such as ground cover, nature objects in urban environment) or to other trajectories (moving objects). Trajectories can enter or can cross a spatial environment, or trajectories can under a spatial environment; they intersect, meet and can be near or far to other trajectories.

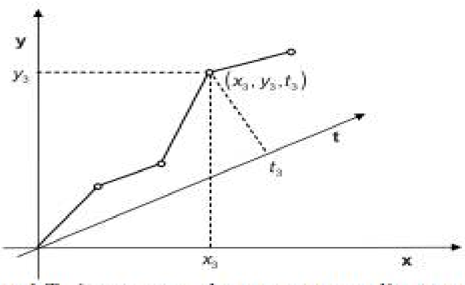

Figure 1 Trajectory - x and y represent coordinates and t represents time

Deploying DKNN and discuss how to run the DKNN algorithm on S4 to process k-NN queries. The DKNN algorithm can be decomposed into several steps, and each step can be handled by a particular type of PEs. We call the resulting scheme DKNNS4. The advantage of DKNN is that even though the master node does not store the positions of objects, it can still determine the search space that contains the k-NNs in just two steps, by first directly determining the candidate strips using the DCS algorithm, and then identifying the final set of strips to search by computing the circle Cq. This is highly beneficial when the algorithm is running in a distributed system. In contrast, most existing search algorithms do not have this property. With those algorithms, the master cannot determine the final region for k-NN search without involving an uncertain number of rounds of communication between the master and slaves, incurring significant communication costs.

1. Survey of Nearest Neighbor Condensing Techniques [1]

In this paper we tried to present and compare sets reduction techniques based on the principle of nearest neighbour. These techniques are of type "condensing". Both techniques are efficient than basic KNN. These techniques have been introduced by the authors to reduce the training set to increase speed and space efficiencies. To ensure the minimalist of this training set we presented some recent proposals using meta heuristics to check the optimality of the resulting set of some KNN reduction techniques. Note that each technique is very effective in a particular area and in special circumstances.

2. Improved Adaptive K NearestNeighbour algorithm using MapReduce [2]

Improve performance of classification algorithm using hadoop tool for large set of data. The accuracy of the algorithm is increase and reduces the time required for execution. In future Multi-threading K Nearest Neighbour algorithm results obtained from cloud platform with hadoop framework using map reduce programme to understand the effectiveness. In the future work, we will further implement other classification algorithms and conduct the experiments and consummate the parallel algorithms to improve usage efficiency and accuracy of computing resources and reduce the execution time.

3. Brute-Force k-Nearest Neighbors Search on the GPU [3]

Finding ways to use highly-optimized GPU library functions is an effective way to achieve both speed and robustness in this important application. Our algorithm advances the state of the art for all but the smallest values of k. It is unique

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Vol. 4, Issue 1, January 2016

in its ability to handle large values of k, and large input datasets. The performance of our algorithm for very small values of k is limited mainly by the performance of the selection step. Possibly this could be improved by allowing one thread block to perform multiple truncated merge sorts in parallel .The disadvantage of this method would be that it complicates the kernel. Approximate kNN search approaches where nearest-neighbour candidates are altered so that not all squared distances need to be computed could benefit from using our truncated merge sort to select the true nearest neighbours from the candidates. This is true for Locality Sensitive Hashing as well as for heuristic approaches.

4. Approximate K-Nearest Neighbor Based Spatial Clustering Using K-D Tree [4]

The great contribution achieved by this research is the use of the approximate k- nearest Neighbor with k-d Tree data structure for spatial clustering, and equate its performance to the brute-force approach. The results of the work performed in this techniques revealed better performance using the k-d Tree, compared to the brute-force approach. The efficiency of the data structure depends on a particular implementation and data set.

Processing Multiple k Nearest Neighbor Queries, introduced a technique that can quickly identify the related queries of a given query q. After finding the related queries, algorithm uses the information to infer a search region that covers the kNN results of q. R-Grid to reduce the distance computations when traversing an R-tree.

5. KNN based Machine Learning Approach for Text and Document Mining [5]

We conclude that KNN shows the maximum accuracy as compared to the techniques Naive Bayes and Term-Graph. The disadvantage of KNN is that its time complexity is high but gives a better accuracy than other techniques. We referred Term-Graph with other methods rather than the traditional Term-Graph. This combination shows a better result than the traditional combination. Finally we referred an information retrieval application using Vector Space Model that give the result of the query entered by the client by showing the relevant document. We will give importance more in future on Reducing Complexity, Increasing Accuracy and Text Summarization.

6.Processing Multiple k nearest Neighbor Queries [6]

In this paper, we introduced a technique that can quickly identify the related queries of a given query q. After finding the related queries of q, our algorithm uses the information to infer a search region that covers the kNN results of q. We employ an R-Grid to reduce the distance computations when traversing an R-tree. The performance results show that the techniques proposed in the paper can speed up the performance and reduce the disk I/O cost of kNN searching.

III. DRAWBACKS OF PREVIOUS SYSTEM

Disadvantage of related work is in most cases considering only location and time as attributes of moving object's movement. The adapting model with embedding knowledge about geospatial conditions (e.g. type of a habitat, climate) that pertain to the location and temporal conditions (e.g. time of the day) leads to a more exact model about moving objects behavior. The most of additional attributes can be extracted from coordinates and time attributes. Knowing space coordinates, attributes such as vegetation, altitude or type of road can be scanned from geospatial maps. Similarly, knowing time attribute, attributes such as season, temperature or rainfall can be get from historical data collected in weather station. The existing techniques to this problem are designed for the centralized setting where query processing takes place on a single server; it is difficult, for distributed setting to handle the high volume of data and concurrent queries that are common in these applications.

IV. PROPOSED SYSTEM MECHANISM

The large volume of data and heavy query workloads call for new scalable solutions. We propose DSI, a distributed strip index, and DKNN, a distributed k-NN search algorithm, to address this challenge. Both DSI and DKNN are designed with distributed processing in mind, and can be easily deployed to a distributed system. DSI is a data partitioning index and is able to adapt to different data distributions. Based on DSI, we present the DKNN algorithm that can directly determine a region that is guaranteed to contain the k-NNs for a given query with only two iterations. This has a clear cost benefit when compared with existing approaches, such as grid-based methods, which require an uncertain number of iterations. We show how the proposed index and algorithm can be implemented with S4. Extensive experiments confirm the superiority of the proposed method. We would like to explore how to evaluate continuous k-NN queries over moving objects using the strip index. For a given k-NN query q, it is very possible that its result (a list of objects) remains relatively stable when objects move with reasonable velocities.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Vol. 4, Issue 1, January 2016

V. SYSTEM ARCHITECTURE

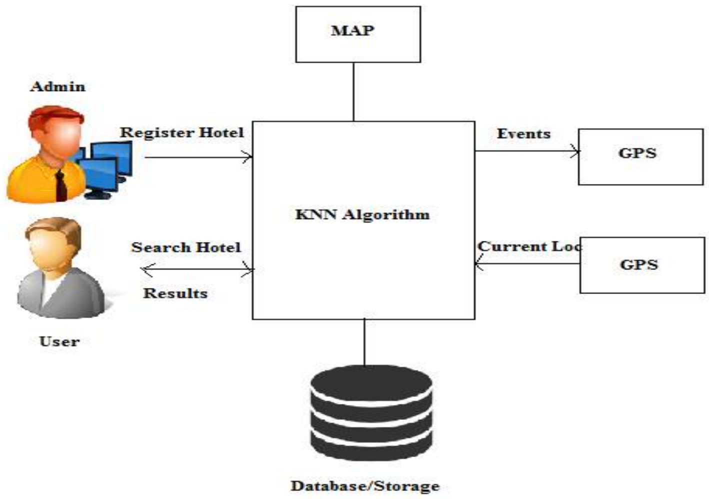

FIG NO. 1 SYSTEM ARCHITECTURE OF PROPOSED SUSTEM.

V. CONCLUSION

In this paper, we introduced a technique that can quickly identify the related queries of a given query q. After finding the related queries of q, our algorithm uses the information to infer a search region that covers the kNN results of q. We employ an R-Grid to reduce the distance computations when traversing an R-tree. The performance results show that the techniques proposed in the paper can speed up the performance and reduce the disk I/O cost of kNN searching. Currently, our algorithm can only deal with kNN queries. In the future, we plan to extend our work to permit various spatial queries (e.g., reverse kNN, aggregate NN and range queries). We plan to explore this subject further in the future.

REFERENCES

[1] MILOUD-AOUIDATE Amal (corres.Author) Computer Sciences Department University of Sciences and Technology HouariBoumediene, USTHB, Algiers, ALGERIA"*Survey of Nearest Neighbor Condensing Techniques*".
[2] Apexa B. Kamdar, Ishan K. Rajani"*Improved Adaptive K Nearest Neighbor algorithm using MapReduce"*.
[3] Shengren Li and Nina AmentaUniversity of California, Davis "*Brute-Force k-Nearest Neighbors Search on the GPU"*.
[4] Dr. Mohammed OtairDepartment of Computer Information Systems, Amman Arab University, Amman, Jordan"*Approximate K-Nearest Neighbor Based Spatial Clustering Using K-D Tree"*.
[5]Vishwanath Bijalwan1, Vinay Kumar2, Pinki Kumari3 and Jordan Pascual4 *1Asst. Prof., Institute of technology Gopeshwar, Chamoli, Uttarakhand, India 3Bansathali University, Rajasthan, India4 Department of Computer Science, University of Oviedo, Spain*"KNN based Machine Learning Approach for Text and Document Mining".
[6]I-Fang Su, Yu-Chi Chung, Pei-Chi Liu, and Chiang Lee "Processing Multiple k Nearest Neighbor Queries".