MATRIX ACADEMIC INTERNATIONAL ONLINE JOURNAL OF ENGINEERING AND TECHNOLOGY
ISSN: 2348-3326,
www.maioj.org © MAIOJET
Volume 3, Issue 2

An Interactive Spatial Data Mining Framework Using Query Pattern Analysis

Mohd Haseeb Ahmad
Asst.Professor , Shri Ramswaroop Memorial College Of Engineering And Management, Lucknow, India.
btechfirstyearc@gmail.com

Vinodini Katiyar
Professor, Shri Ramswaroop Memorial University, Lucknow, India.
drvinodini@gmail.com

S.Q.Abbas
Professor & Director, Ambalika Institute of Management And Technology, Lucknow, India.
qrat_abbas@yahoo.com

Abstract- Spatial data sets always draw attention towards it because of hefty section of social and informal needs are associated with it. As size and attributes of spatial data differ rigorously from traditional data, spatial data mining becomes quite tedious and typical, as size and usages of spatial data is increasing now a days quite noticeably, hence a proper procedure to mine becomes more important. We suggest an interactive frame work to mine spatial data sets which will produce better response timed results when query gets repeated.

Keywords-spatial data mining, spatial database, data mining, Interactive approach, analyzer program, data shelf.

I. INTRODUCTION

SPATIAL data mining, i.e., discovery of encouraging features and designs that may indirectly exist in spatial databases in many applications, might become the main feature of interest. There may be several cases where we want to this division of mining in role as:

1. How to use the end user as an trigger firing agent to the spatial data set ?considering:
   - a farmer's location at time t makes him concerned about topographical description of the location, the productiveness of the soil, the climate estimate etc. or
   - a politician's location in a particular time interval makes him concerned about the diversity of population, religion, sects or
   - Armour deployment in some area focussing via satellite images or
   - Applications needing care for a large number of geographically dispersed mobile users collaborating on a domain etc.
2. How to produce the pictorial interactions to the user?
   - An airplane's position makes it is alarmed about the air pressure calculation, pollution percentage, altitude etc.
3. How to include user choices to the interface?
   - It is necessary to cope up with the user demands as user wants his required query executed up to maximum possible accuracy.

II. RELATED WORK

SPATIAL DATA MINING

A. The Definition of Spatial Data Mining

Spatial data mining (SDM) revenues finding out wisdom from spatial database, In general, the Data Mining process includes three basic steps:

1) Focussing: This step includes tracing the data to be mined, choosing appropriate tools required in extracting, deploying, and probing data from a large data warehouse.

2) Uncovering: This step includes using one or more techniques to extract patterns of attention, diminishing data errors and deceived searches.

3) Authentication: The mined information is assessed by means defined by the mining criteria.

B. The Characteristics of Spatial Data

Spatial data are generally categorized as they are unrestricted and commonly huge, possess Altitudinal reliance (the framework has a significant influence) and heterogeneity or nonstationarity.

Thus how to recover vanished data and guesstimates for the fundamental distribution of data develops major complications in mining of data tree.

Following are the common attributes of spatial data comparing data with normal data:

- Spatial data is regularly very bulky in size whereas traditional data is not.
- Spatial data is elastic or unfixed because shape and content describing reference frame has an impact on it where as traditional data is not open ended.
- Spatial data has some kind of effect or addiction on another spatial data in neighboring location where as traditional data does not have influences.
- Spatial data is diverse or non-stationary as it may vary over locations or time where as traditional data is not dissimilar.
- Spatial data is rapidly changed over locations as in the case of remote sensing where as traditional data does not change rapidly.
- Spatial data holds property of insufficiency as it may be lost or not acquired due to some constraints where as traditional data is said to be inconsistent if it is insufficient.

III. THE TYPICAL METHODS OF SPATIAL DATA MINING

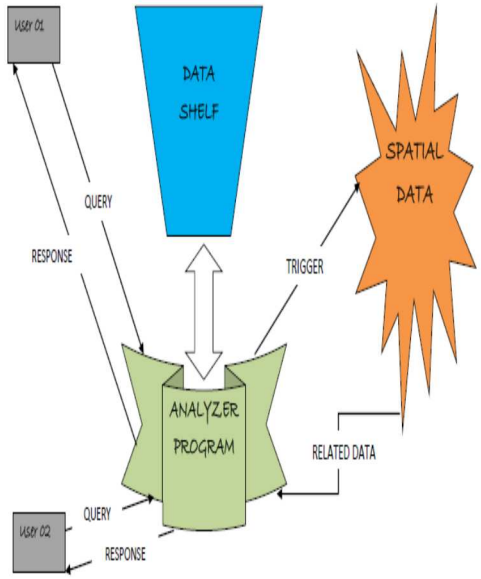
There have been many researches in this area during recent years [1-7] .Some of them are listed below with the technique of mining:

- A research customs a game theoretic framework to suggest the fair value of information shared between two parties. Both the parties mine the information and then consequence to cost efficiency and deliver a means for attaining decision rules for pricing shared information. [2]
- A research provide a transaction-free approach to mine colocation patterns by using the concept of proximity neighbourhood. The colocation pattern detection progression finds the subcategories of structures which are recurrently located organised and this algorithm embraces a novel multiresolution pruning technique. [3]
- A research propose and analyse two novel algorithms for outlying sensor identification and incident Boundary detection and targets the identification of remote sensors and the discovery of the influence of events in radar networks. [4]
- A research present a graph-based style for mining geospatial statistics. The system uses error-tolerant graph matching to find communications between the detected image geographies and the geospatial trajectory data. Diagram matching is used as a query instrument to reply the spatial query. [5]
- A research propose a well-organized method to develop association rules from spatial statistics using Peano Count Tree (P-tree) assembly. P-tree assembly offers a lossless and compressed representation of spatial data. Based on P-trees, a competent association rule mining algorithm PARM with fast support calculation and substantial pruning procedures is presented to improve the effectiveness of mining process. [6]
- A research suggest a method for spatial generalization and aggregation of movement data, which renovates routes into shared flows between zones assuming that no predefined areas are in focus and devise a method of partitioning the zone into appropriate areas.[7]

IV. OUR APPROACH FOR INTERACTIVE FRAMEWORK

The proposed protocol works as follows:

- User initiates the query process replied by analyzer program of framework.
- Analyzer program selects the appropriate mechanism for data fetch.
- The analyzer program then triggers to the spatial data set using the appropriate selection.
- The related data is then captured by analyzer program.
- The data shelf is then mirrored about the captured response.
- The analyzer program is then responded for the user queries.



Pros

- Interaction with user in an efficient manner.
- If hit at data shelf is high then workload will be low then performance is high.

Cons

- The frame work has to come up with analyzer program and shelf maintenance.
- Whatever the fetch required the trigger is to be directly written at runtime.
- Shelf size is supposed to be small.

V. EXPERIMENTAL RESULTS

DATA SHELF

Query	Response set
Q1	Đ4,µ1,n1
Q2	Đ5,µ2,n2
Q3	Đ2,µ3,n3
Q2	Đ5,µ2,n2/n2,Q2
Q5	Đ4,µ1,n1
Q6	Đ4,µ5,n5
Q2	Đ5,µ2,n2/n2,Q2(Fr)

ANALYZER PROGRAM

query	Data triggers set	Time slices set	selection	To data shelf
Q1	Đ1 Đ2 Đ3 Đ4	µ1 µ2 µ3 µ4	n1	Đ4,µ1,n1
Q2	Đ1 Đ2 Đ3 Đ5	µ1 µ2 µ3 µ5	n2	Đ5,µ2,n2
Q3	Đ5 Đ2 Đ3 Đ6	µ5 µ2 µ3 µ6	n3	Đ2,µ3,n3
Q4	Đ7 Đ4 Đ5 Đ2	µ7 µ4 µ5 µ2	n4	Đ4,µ4,n4
Q5	Đ1 Đ2 Đ3	µ1 µ2 µ3	n1	Đ4,µ1,n1
Q6	Đ2 Đ5 Đ4	µ2 µ5 µ4	n5	Đ4,µ5,n5
Q7	Đ1 Đ2 Q2	µ1 µ2	n1	Q2,µ1,n1

VI. CONCLUSION

Given at the set of triggers (Đ1 Đ2 Đ3 Đ4 Đ5 Đ6...) having time slices of (µ1 µ2 µ3 µ4 µ5 µ6…), if query Q1 may be targeted by set (Đ1 Đ2 Đ3 Đ4) with slices (µ1 µ2 µ3 µ4) selection n1 is made and data shelf is mirrored by Đ4, µ1, n1.

Query Q2 may be targeted by set (Đ1 Đ2 Đ3 Đ5) with slices (µ1 µ2 µ3 µ5) selection n2 is made and data shelf is mirrored by Đ5, µ2, n2.

Query Q3 may be targeted by set (Đ5 Đ2 Đ3 Đ6) with slices (µ5 µ2 µ3 µ6) selection n3 is made and data shelf is mirrored by Đ2, µ3, n3.

Query Q4 may be targeted by set (Đ7 Đ4 Đ5 Đ2) with slices (µ7 µ4 µ5 µ2) selection n4 is made and data shelf is mirrored by Đ4, µ4, n4.

Query Q5 may be targeted by set (Đ1 Đ2 Đ3) with slices (µ1 µ2 µ3) selection n1 is made and data shelf is mirrored by Đ4, µ1, n1.

Query Q6 may be targeted by set (Đ2 Đ5 Đ4) with slices (µ2 µ5 µ4) selection n5 is made and data shelf is mirrored by Đ4, µ5, n5.

Query Q7 may be targeted by set (Đ1 Đ2 Q2) with slices (µ1 µ2) selection µ1 is done keeping tracked by Q2,µ1 on behalf of q2 repeat and selection n1 is made and data shelf is mirrored by q2,µ1,n1 noting that this query is frequent(Fr).

Thus we may conclude that next two references to Q2 will have shorter response time.

VII. ACKNOWLEDGMENT

This work is supported by Shri Ramswaroop Memorial University (SRMU) Lucknow and Shri Ramswaroop College of Engineering and Management (SRMCEM) Lucknow.

REFERENCES:

[1] Vinodini Katiyar; Mohd Haseeb Ahmad; S.Q.Abbas "Novel Approach for Interactive Spatial Data Mining", MATRIX Academic International Online Journal of Engineering and Technology, ISSN: 2348–3326 © MAIOJET. Volume III Issue I, April 2015.

[2] Saygin, Y.; Reisman, A.; YunTong Wang, "Value of information gained from data mining in the context of information sharing.", IEEE Transactions on Engineering Management, Volume: 51, Issue: 4, 441 – 450, 2004 .

[3] Yan Huang; Shashi Shekhar; Hui Xiong, "Discovering colocation patterns from spatial data sets: a general approach", IEEE Transactions on Knowledge and Data Engineering, Volume: 16, Issue: 12, 1472 – 1485, 2004.

[4] Weili Wu; Xiuzhen Cheng; Ding, M.; Kai Xing; Fang Liu; Ping Deng , "Localized Outlying and Boundary Data Detection in Sensor Networks", IEEE Transactions on Knowledge and Data Engineering, Volume: 19, Issue: 8, 1145 – 1157, 2007.

[5] Sidharta Gautama; Bellens, R.; De Tre, G.; Philips, W, "Relevance Criteria for Spatial Information Retrieval Using Error-Tolerant Graph Matching", IEEE Transactions on Geoscience and Remote Sensing, Volume: 45, Issue: 4, 810 – 817, 2007.

[6] Qin Ding ; Qiang Ding ; Perrizo, W., "PARM—An Efficient Algorithm to Mine Association Rules From Spatial Data", IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, Volume: 38, Issue: 6, 1513 – 1524, 2008.

[7] Adrienko, N.; Adrienko, G., "Spatial Generalization and Aggregation of Massive Movement Data", IEEE Transactions on Visualization and Computer Graphics, Volume: 17, Issue: 2, 205 – 219, 2011 .