DESIGN OF SPATIO---TEMPORAL PATTERNS AS A QUERY MECHANISM USING DATA MINING

Md Moizuddin*

Imran Hassan*

Mohammed Younus**

Khazi Mohammed Farooq**

Abstract—

The process of discovering meaningful new correlations, patterns and trends by sifting through large amounts of data stored in repositories and by using pattern recognition technologies as well as statistical and mathematical techniques. The exploration and analysis of large quantities of date in order to discover meaningful patterns and rules. The nontrivial extraction of implicit, previously unknown, and potentially useful information from data Our world is changing at an increasing pace causing the validity intervals of information to shrink. Therefore the change of information becomes important information itself—an important information resource that should be exploitable by query language. Beyond data mining that often serves the purpose of identifying interesting objects in special or spatiotemporal data sets, queries are of interest that find our about relationships between objects once they have been identified. In the realm of spatiotemporal database this means to find out, in particular, about the changes the objects and their relationships undergo. Moreover many applications require not only the detection of one single change, but rather look for sequences of changes describing particular developments.

Keywords: Data Mining, Spatio-Temporal Mining ,Spatio Temporal Patterns

* COLLEGE OF COMPUTER SCIENCE, KING KHALID UNIVERSITY

** COLLEGE OF COMPUTER SCIENCE AND INFORMATION SCIENCES, KING SAUD UNIVERSITY

INTRODUCTION

We propose the concept of spatiotemporal patterns as a systematic and scalable concept to query developments of objects and their relationships. Based on our previous work on spatiotemporal predicates, we outline the design of spatiotemporal patterns as a query mechanism to characterize complex objects behaviors in space and time.

Spatiotemporal patterns can be applied in may different areas of science, for example in geosciences, geophysics, meteorology, ecology and environmental studies, study of climatic changes.

EXISTING SYSTEM

Existing spatiotemporal database systems and query languages offer only basic support to query changes of data. Most of these systems allow the formulation of queries that ask for changes at particular time points. It is often very difficult changes. In other words, existing query languages do not offer a systematic, scalable concept to query developments of objects and relationships; instead they require the user to encode these by a number of individual conditions.

It also requires the formulation of additional side conditions. Moreover, this approach does not work in queries for arbitrary numbers of changes.

Spatiotemporal patterns enable the formulation of queries about complex object developments their integration into query languages allow the formulation of queries that are currently not possible. Moreover, spatiotemporal patterns can also simplify the formulation of queries that are possible with existing languages.

PROPOSED SYSTEM

The intervals for which certain pieces of information remain valid become shorter and shorter. Therefore, the change of information becomes important information itself an important information resource that should be exploitable by query languages. In the area of temporal data bases this issue has been addressed for quite some time. Availability o spatio. Temporal Data.

- Meteorology: all kinds of weather data, moving storms, tornados, developments of high pressure areas , movement of precipitation areas, changes in freezing level, droughts.
- Biology: animal movements, mating behavior, species relocation, and extinction.

- Crop Sciences: seasonal grasshopper infestation, harvesting, soil quality changes, land usage management……
- Forestry: forest growth, forest fires, hydrology patterns, canopy development, planting tree cutting, planting tree planting……
- Medicine: patients cancer developments, supervising developments in embryology…
- Geophysics earthquake histories, volcanic activities and prediction.

Many Human-Related Activities are Related to Changes in Spatial Information as well

- People: Movements of terrorists/criminals/spies, movement of people in emergency situations, pedestrian patterns/habits…
- Cars/trucks/taxis : tracking, rerouting, fleet management,
- Urban planning, parcel management, development of social areas, urban
- Economics, tourism planning, bus routes,….
- Crime/disaster prevention : Risk area analyses, resource allocations (police, health care, fire stations
- History : Country expansions, reunifications , tribe movements,
- Military : missile tracking, troop movements,
- Planes/ships : routes, detours,

Data Mining

Data mining is a system of searching through large amounts of data for patterns.

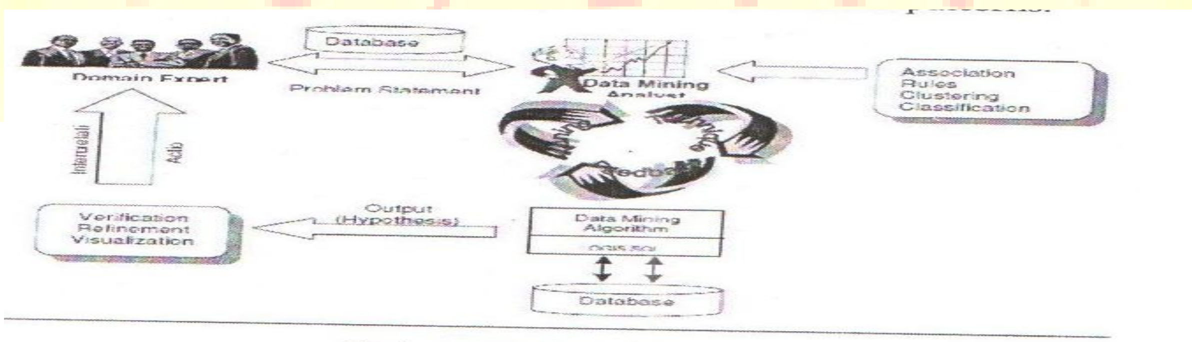


Fig. 1: A Data Mining Process

Spatio - Temporal Patterns

Regular structures in space and time, in particular, repeating structures, are often called patterns. Patterns that describe changes in space and time are referred to as spatiotemporal patters.

One important characteristic of change in spatial data that makes it interesting for data processing by scientists and others; interesting phenomena are those that are not random but rather follow certain rules. In other words, applications are interested in spatial data changes that exhibit a certain regular structure.

Limitation of Existing System

One goal of spatiotemporal database research is to enable the intelligent use of the collected data, a main problem is that of extracting relevant and useful information from the data.

There are two areas of database research addresses two different aspects of the problem

- Data mining which is also sometimes referred to as knowledge discovery is concerned with the discovery of patterns in large data sets.. In particular, spatiotemporal data mining has the goal of identifying interesting structures and patterns in spatiotemporal data sets.
- Query languages work on data that is well structured and that is stored in rather fixed formats obeying schema definitions and possibly additional integrity constraints. The distinctive feature of query languages (compared to
- Data mining) is the assumption that the knowledge in a database is already present, whereas the goal of data mining is to discover it. Query languages are used to find out further relationships among objects.

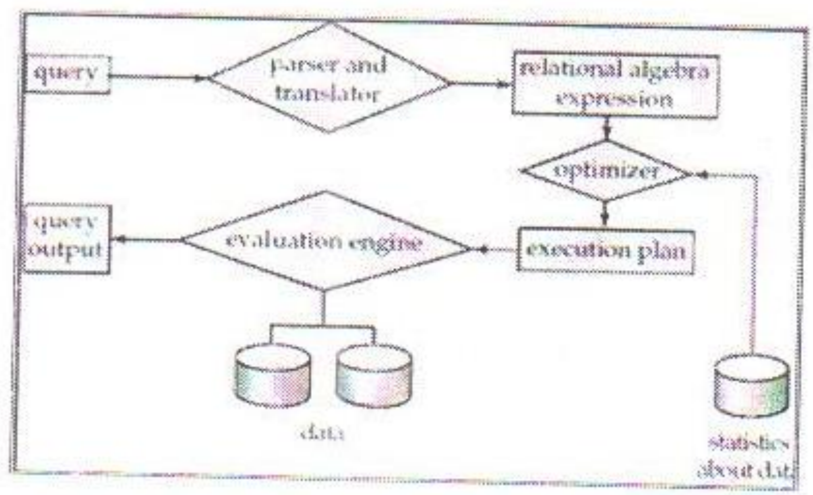
Design of Spatiotemporal Patterns as a Query Mechanism

Spatiotemporal predicates allow the formulation of many queries in a rather direct way of expressing spatiotemporal relationships directly on attributes of the involved objects.

- How can the notion of spatiotemporal predicates be generalized without compromising their current useful applications ?
- What is the suitable query language to specify aggregations over spatiotemporal predicate results?
- How the earlier version different from the proposed one?

- Is a spatiotemporal pattern providing a way to describe hierarchical or nested spatiotemporal predicates?
- Whether we can use recursive grammar or regular expression?

General Query Processing



Parsing and Translation

Translate the query into its internal form. This is then translated into relational algebra.

- Parser checks syntax , verifies relations

EVALUATION

- The query – execution engine takes a query –evaluation plan, executes that plan, and returns the answers to the query.
- A relational algebra expression may have many equivalent expressions
- For example

$$\sigma_{balance<2500}(\prod_{balance}(account)) \text{ is equivalent to } \prod_{balance}(\sigma_{balance<2500}(account))$$

Spatio-Temporal Approach

We have introduced a concept of spatiotemporal patterns that can be employed in query languages to express queries about the development of spatiotemporal objects and their relationships. To help copying with the complexity, we believe that providing effective access to spatiotemporal data through expressive query languages is an important step to help scientists and ordinary users to exploit the information that is contained in fast growing spatiotemporal data sets. Of the query language we also sketched the design of a visual query language that makes the expressiveness of spatiotemporal patterns accessible to a large group of users.

Research need in response to the problem is to develop mechanisms to test and validate spatiotemporal data mining results, particularly that test the validity of spatial and temporal relations, and to reconcile discrepancies in data.

SCALABLE INCREMENTAL PROCESSING OF CONTINUOUS SPATIO-TEMPORAL QUERIES

Hence we are presenting a scalable and incremental framework for continuously evaluating continuous spatio-temporal queries. The scalability is achieved by employing a shared execution paradigm for continuous spatio-temporal queries. With the shared execution, queries are indexed in the same way as data. Thus, evaluating a set of concurrent continuous spatio-temporal queries is reduced to a join between a set of moving objects and a set of moving queries. Incremental evaluation is achieved through computing only the updates to the previously reported answer/

We have Three Classifications of Queries

- **Stationery Queries**
- **Moving Queries**
- **Predictive Queries**

VISUAL QUERY MECHANISM

Spatiotemporal predicates allow the formulation of many queries in a rather direct way by expressing spatiotemporal relationships directly on attributes of the involved objects. The more advanced queries that require spatiotemporal patterns quickly tend to become quite complex. Moreover, the fast growing set of available spatial and spatiotemporal data and its wide dissemination through the internet increases the class of possible users for these data. Offering ordinary users access to spatiotemporal data is therefore becoming a more and more important issue that can be addressed by developing a visual query language and a corresponding user interface. In particular, it is important to make data base accessible to users without formal training in databases and query languages, such as scientists, administrators or other end users. In many cases, these users do not have the time or are not willing to learn formal query language.

From this point of view, even spatiotemporal predicates embedded into a query language like STQL.

A visual notation can express relationships in many cases implicitly where textual notations require the explicit application of operations and predicates.

For Example

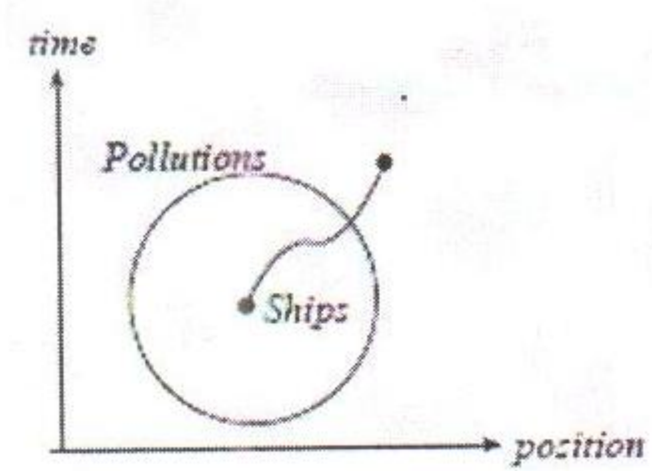


Fig. 1.2: Visual Specification of the Leaves Predicate

The picture describes the following query and can be automatically translated into it.

SELECT sname, pname

FROM shipsf, Pollutions

WHERE Pos (Inside meet Disjoint) Reg

To extend the developed visual query notation (and the user interface) to cope with more general forms of spatiotemporal predicates and patterns, a number of issues have to be addressed. A refinement of the visual query in Figure 1.2 that is obtained by adding a constraint on the time the ship needed to leave the pollution region. We can express this by projecting two points onto the time axis and adding a condition to it. The resulting visual query asks only for those ships that have left the region within 30 minutes, see Figure 1.3.

This visual query will then be translated into the following STQL query.

SELECT sname, pname

FROM Ships, Pollutions

WHERE Pos(For(<30 min, Inside) meet Disjoint) Reg

Two other useful extensions are to express numerical spatiotemporal predicates and to allow for hierarchical specifications. To use numerical spatiotemporal we need additional visual elements to refer to numerical spatiotemporal operations like Distance to be able to express numerical constraints on developments. In order to be able to express general spatiotemporal patterns that require nested sequence conditions we need a mechanism to refer to visual specifications from within other specifications.

CONCLUSION & DISCUSSIONS

Existing spatiotemporal data models and query languages offer only basic support to query changes of data, in particular, although these systems often allow the formulation of queries that ask for changes at particular time points, they fall short of expressing queries for sequences of such changes.

We have introduced a concept of spatiotemporal patterns that can be employed in query languages to express queries about the development of spatiotemporal objects and their relationships. To help copying with the complexity of the query language we also have sketched the design of a visual query language that makes the expressiveness of spatiotemporal patterns accessible to a large group of users.

We believe that providing effective access to spatiotemporal data through expressive query languages is an important step to help scientists and ordinary users to exploit the information that is contained in fast growing spatiotemporal data sets.

FUTURE WORK

Spatiotemporal patterns can be applied in many different areas of science, for example, in geosciences, geophysics, meteorology, ecology and environmental studies. Since users in the in these areas typically do not have extended formal computer training. It is often difficult for them to use advanced query languages. A visual notation for spatiotemporal patterns can help solving this problem. In particular, since spatial objects and their relationships have a natural graphical representation, a visual notation can express relationships in many cases implicitly where textual notations require the explicit application of operations and predicates. Future work on the visualization of spatiotemporal predicates, we will sketch the design of visual language to formulate spatiotemporal patterns.

REFERENCES

(1) Jiawei Han , Michline Kamber, "Data Mining Concepts and Techniques" , Morgan Kaufmann Publishers, 2002.

(2) Abraham Siberschatz, Hentry F. Korth and S.Sudharssan, "Database System Concepts", 4th edition, Tata McGraw Hill 2002.

(3) Raghu Ramakrishnan & Johannesgerhrke, "Data Base Management System" Mc Graw Hill International Edition 2000

(4) G.Ariyav An Overview of TQuel .ACM Transactions on Database Systems , 11(4):499-527,1986

(5) In A.U.Tansel.J.Clifford S.Gadia S.Jajodia A.Segev, and R Snodgrass , editors, Temporal Databases : Theory Design and implementation, pages 248-270 , The Benjamin /Cummings Publishing Company, 1993.

(6) R.T.Snodgrass, editor, The TSQL2 Temporal Query Language Kluwer Academic Publishers, Boston, MA, 1995

(7) S.Imfeld Time, Points and Space - Towards a Better Analysis of Wildlife Data in GIS, Dissertation , University of Z'urich, 2000.

(8) D.J.Hand and H.Mannila and P.Smyth. Principles of Data Mining , MIT Press, Cambridge, MA, 2001

(9) G.M.Kuper, L.Libkin, J.Paredaens, editor, Constraint Databases, Springer –Verlag, Berlin, 2000

(10) Butterfield B.Gahegan, M.Miller, H.Yuan, M.(2001), Geospatial data mining and knowledge Discovery UCGIS white paper on Emergent Research Theames.