OPPORTUNITIES FOR KNOWLEDGE DISCOVERY IN SPATIO-TEMPORAL INFORMATION SYSTEMS

Tamas Abraham and John F. Roddick
Advanced Computing Research Centre
School of Computer and Information Science
University of South Australia
The Levels SA 5095
Email:{abraham, roddick}@cis.unisa.edu.au

ABSTRACT

Spatial Information Systems and their recent temporal extensions typically store large volumes of geo-referenced information. Having such size, it becomes increasingly difficult to explore their contents with current querying techniques. In this paper, we examine how data mining methods can help users in the analysis of the contents of Spatial and Spatio-Temporal Information Systems. We review existing spatial applications and investigate how they can be extended to deal with time. We also look at new, alternative methods that utilise the inherent structure of spatio-temporal information as well as its rich semantics to derive rules about changes and movement.

Keywords

Knowledge Discovery, Data Mining, Spatial Information Systems, Spatio-Temporal Data

INTRODUCTION

Information stored in digital knowledge repositories has never been so abundant. With the rapid advance in hardware technologies and the increased affordability of computers, more organisations are turning towards storing increasing amounts of data in their information systems. Spatial databases, and their temporal extensions, are examples of the technology being used to support this. The benefits of the availability of detailed information can be seen in *inter alia* improved decision support, database integrity enforcement and query optimisation. The exploration of these huge data sets with existing querying techniques, however, remains a burden. To alleviate this problem and offer additional tools for analysis, *data mining* and *knowledge discovery in databases* provide techniques to extract useful, implicit information from large databases, combining methods from artificial intelligence, statistics and database research (Fayyad, et al. 1996b). Some of these techniques have been successfully used in applications dealing with transactional and relational data. On a lesser scale, spatial and temporal applications are also in existence (Fayyad, et al. 1996a, Koperski, et al. 1996).

The comparatively small number of data mining techniques available for spatial information systems can be explained by not only the more widespread use of other data types, but the complexity of spatial and spatio-temporal data vs. relational data. Spatial Information Systems (SIS) deal with information about our surroundings, be that in the form of maps in Geographic Information Systems (GIS), aerial or satellite images in Image Processing systems, land parcels in Land Information Systems (LIS), or networks such as the telephone lines or power grids in Automated Mapping/Facilities Management Systems (AM/FM) (Abel 1996). They store spatial in addition to aspatial data, often using proprietary representational techniques, commonly based on the two best known alternatives, the location-oriented raster or the spatial entity-oriented vector structure, which are based respectively on the field and object data models. For overviews on the latest developments and advances in spatial database research, refer to (Abel 1996, Guenther & Buchmann 1990, Güting 1994, Medeiros & Pires 1994, Paredaens 1995).

With the introduction of the temporal dimension, spatio-temporal information systems can reason effectively on a variety of new problems, ranging from environmental monitoring and impact assessment, resource management, decision support, administration, real-time navigation, and transportation scheduling to data quality and integrity enforcement. This is achieved by the incorporation of a number of new features that surpass current SIS capabilities. In particular, enhanced analytical power, process visualisation and the integration of multiple data sources are available (Abraham & Roddick 1996, Langran 1992). Analysis and decision support can be further aided by the inclusion of data mining tools.

The rest of this paper is organised as follows. In the next section, we revisit briefly the achievements and progress in data mining in Spatial Information Systems and discuss problems associated with exploring spatial data. In the following section, we investigate the application of current and new data mining techniques to existing spatio-temporal systems. Finally, we present some conclusions and ideas for future research directions.

SPATIAL DATA MINING

Data Mining is an emerging field of computer science that combines elements of artificial intelligence, statistics and database theory to find implicit information in large data sets. A typical discovery process, often guided by some form of background knowledge, uses a search algorithm modified to work on large databases to extract previously unknown patterns with a high level of statistical significance, or "interestingness". Common data mining tasks include deriving the general characteristics of data (e.g. via rule induction), classifying them into different groups (e.g. constructing decision trees based on database contents), extracting association rules between attributes or objects, finding relationships between individual data items (e.g. causal or functional relationships), and detecting trends and deviations (Chen, et al. 1996). Methods that we shall use in later discussions include the attribute oriented induction technique (Cai, et al. 1991) to mine characteristic and discriminant rules, the original association rule extraction algorithm (Agrawal, et al. 1993) and its later extensions, and various clustering techniques, e.g. (Ng & Han 1994). For further information, please refer to the monographs (Fayyad, et al. 1996c, Piatetsky-Shapiro & Frawley 1991), or some of the existing overviews (Han 1996, Rainsford & Roddick 1996). For the more interested readers, general discussions on the future of knowledge discovery also exist (Fayyad, et al. 1996b, Mannila 1997).

Some of the existing data mining techniques have already been applied to Spatial Information Systems. In image databases, pattern matching and analysis have been used to detect volcanos on the surface of Venus (Bell, et al. 1994), or recognise cyclone formation (Stolorz, et al. 1995). Other similar work is overviewed in (Fayyad, et al. 1996a). In geographic information systems, generalisation techniques have been employed to derive the characteristics of spatial and associated attribute data (Lu, et al. 1993). Interactive mining of spatial association rules has also been investigated (Koperski & Han 1995), but most attention so far has been paid to characterising spatial information by finding their distribution patterns using clustering techniques. Clustering, in contrast to generalisation methods such as attribute oriented induction, does not require *a priori* knowledge (commonly conceptual hierarchies, see Figure 1 for examples) to guide the discovery process. Traditional clustering algorithms from statistics, however, prove inefficient for large data sets. Therefore, several improvements and alternatives have been proposed in the literature (Ester, et al. 1995, Ng 1996, Ng & Han 1994, Zhang, et al. 1996, Wang, et al. 1997). For a more detailed overview of achievements in spatial data mining, please refer to (Koperski, et al. 1996).
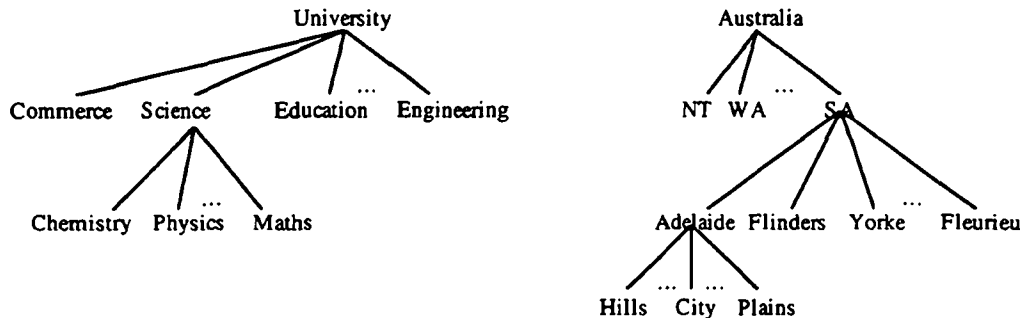

Figure 1: Example Conceptual Hierarchies

In spite of these attempts, few applications are currently available for general use. The aforementioned systems for volcano and cyclone recognition, and some others described in (Fayyad, et al. 1996a), including earthquake detection and star categorisation are specialised systems designed to solve specific (mainly scientific) problems. This contrasts with systems that operate on relational data, where commercially oriented applications dominate. It is not yet clear whether a spatial Knowledge Discovery System (KDS) capable of dealing with a number of spatial discovery problems on several data representational formats will be available in the future or even be a practical object to build. Because of the diversity of SIS, it is foreseeable that each spatial KDS will have to be finetuned to an underlying SIS to some degree. This dependence on the data representation techniques used by the underlying system may prevent the development of a generic spatial knowledge discovery system.

Another problem is that although well-known methods for characterisation and associations can be used with virtually all data types, certain discovery techniques may only apply to specific. Despite these restrictions, the existence of commonly applicable techniques such as the ones mentioned above means that the development of generic spatial data mining *methods* remains a realistic and worthwhile aim. The GeoMiner prototype (Han, et al. 1997) is one such attempt to house a variety of tools in a single system, with spatial data stored using the SAND architecture (Aref & Samet 1991).

Finally, an important issue in data mining that is especially valid for the spatial domain concerns the quality of

data available for knowledge discovery. This has been recognised as having comparable significance to the data mining step in the knowledge discovery process (Mannila 1997). Some companies are finding now that missing or erroneous data may invalidate considerable investments in data warehouses and data mining tools. Noise handling remains a typical area in data mining research (Holsheimer & Siebes 1994, Quinlan 1986). Consequently, SIS may need to be designed to facilitate data mining as opposed to current attempts that try to find ways of extracting new information from pre-compiled, possibly noisy and incomplete spatial and attribute data. Although this complicates the design process and requires some pre-conceived ideas about what will be looked for in the new information system, current data mining techniques will perform better on KD designated SIS and may help the emergence of new techniques for more powerful discoveries.

SPATIO-TEMPORAL DATA MINING

In this section, we first look at some spatio-temporal models and other issues related to the development of Spatio-Temporal Information Systems. Following that, we investigate the kinds of knowledge likely to be sought after with data mining tools, and describe how existing spatial knowledge discovery techniques can be extended to deal with spatio-temporal data. We also discuss two new ways that exclusively utilise the temporal aspects of this data.

Spatio-Temporal Information Systems

Current Spatial Information Systems store an up-to-date, but static record of our environment. They fail to capture its dynamic nature, which considerably limits their reasoning capabilities. To remedy the situation, several attempts have been made to incorporate time into geo-referenced information over the past few decades. Time, location and theme have long been recognised as the building blocks of geographic data (Sinton 1978), but only recently have we seen a variety of spatio-temporal models and working system implementations. Basoglu and Morrison (1978) were among the pioneers in designing a file structure for historical US county boundary data, but it was Armstrong (1988) who first discussed alternative models to primitive archiving scenarios in detail. Among them is the snapshot model which records full database states at progressive time points. This is further improved by a proposal to only store changes from previous versions. Other well-known models include the Space-Time composite model that fragments a base map into smaller and smaller units with their own history as it undergoes change (Langran & Chrisman 1988), a time based method that uses event chains to record the evolution of thematic data for a given geographic area over time (Peuquet & Duan 1995), and a three domain model that represents semantics, space and time as separate, interlinked domains to describe geographic processes and phenomena (Yuan 1996). Alternative proposals found in the literature are models based on relational databases (Rafaat, et al. 1994), the object-oriented paradigm (Bonfatti & Monari 1993, Rojas-Vega & Kemp 1995, Wachowicz & Healey 1994, Worboys 1992), logical versioning (Medeiros & Jomier 1994) and the Temporal Map Set concept (Beller 1991). All these models implement some, but not all, of the requirements of temporal GIS discussed by some authors (Al-Taha & Barrera 1990, Armstrong 1988, Easterfield, et al. 1991, Hazelton 1992, Kemp & Kowalczyk 1994, Roshannejad & Kainz 1995, Yeh & Viémont 1992). Indeed, the creation of a universal spatio-temporal model may not be a realistic goal. Instead, modeling solutions concentrate on capturing change and movement in order to facilitate the answering of certain types of queries. The wider range of queries a system can effectively answer, the more general purpose, and thus more desirable, it can be considered.

In addition to modeling, other database issues need to be addressed before functional spatio-temporal information systems become commercially viable. Abstract spatio-temporal data types and relationship operators must be combined into space efficient data structures with fast indexing methods to satisfy querying requirements. Query languages with both spatial and temporal operators and the underlying mathematical formalisms must be developed, with sophisticated visualisation tools to communicate dynamic processes to users. Existing systems already implement and utilise some of these features. They can be based on available GIS, such as the TEMPEST temporal GIS prototype which employs the model of event chains as mentioned above (Peuquet & Wentz 1994), or Oogeomorph, which is a set of object-oriented classes (Raper 1995). Both solutions access existing, conventional GIS via an appropriate interface while adding temporal capabilities. Fully developed scientific databases, such as GAEA, an object-oriented system to study global change (Hachem, et al. 1992), and QUEST, a prototype built to validate techniques and demonstrate the feasibility of building information systems for atmospheric and earth science databases (Muntz, et al. 1995), go further by having their own spatio-temporal data types, querying and display operators. In addition, temporal extensions of existing tool sets such as the one described by Story & Worboys (1995) can help automate the creation of spatio-temporal applications. For more information, refer to (Langran 1992), which provides a detailed discussion on temporal GIS requirements and modeling techniques. A survey of papers spanning other spatio-temporal issues and new

spatio-temporal model proposals can also be found in (Abraham & Roddick 1996).

Spatio-Temporal Rules

When we consider mining spatio-temporal data, two questions must first be answered: *what* to mine and *where* to mine it, i.e., what kind of information we intend to extract from the spatio-temporal information system and what will the format of this database be?

Although we emphasised that the development of useful techniques are of a high priority, we first briefly discuss the platforms currently available for implementation and testing purposes. Object-oriented databases offer probably the best choice to represent spatio-temporal information: they convey its hierarchical nature and allow histories of objects to be recorded within the objects. However, efficiently querying object-oriented databases remains problematic, and they have not yet gained the general acceptance relational databases enjoy. Relational databases, on the other hand, are often coupled with proprietary representations in spatial information systems, which makes the development of discovery algorithms difficult. Some of the architectures preferred by research groups also have limitations. The SAND architecture used to mine spatial characteristic and association rules provides operators to handle spatial data, but it does not directly handle time. Temporal extensions of the public domain GRASS GIS (Neteler 1997) are available, but at present purely relational spatio-temporal models could still be the best choice for a mining platform. Rafaat, et al. (1994) and Beller (1991) both proposed architectures that enable the use of existing knowledge discovery algorithms with only minor modifications on vector and grid-based representations, respectively.

After selecting an appropriate platform, we must determine the kind of information we intend to look for in the data. The availability of temporal information introduces the possibility of mining for new types of rules that complement existing ones. In particular, we are interested in capturing changes and movement. We shall first discuss two new types of rules that we can use for this purpose, then describe extensions to spatial data mining methods that also deal with temporal information.

(a) Meta-rules

In our definition, meta-rules are rules derived from comparing existing rule sets that were generated at different times on the same domain. Thus, *spatio-temporal meta-rules* describe changes in rule sets obtained from consecutive spatial snapshots of the data in the spatio-temporal information system. In this way, meta-rules can characterise the evolution or constancy of existing patterns, which is useful in predicting future trends in the modelled environment. Survey data is one example for information that becomes available after regular collections. Each data set can be mined for patterns to yield possibly useful rules about the environment at collection times. Meta-rules can then enhance analysis by providing further insight into the changes between surveys. Figure 2 illustrates how spatio-temporal data can be mined at selected time points that yield snapshot rule sets which are then used to produce meta-rules.
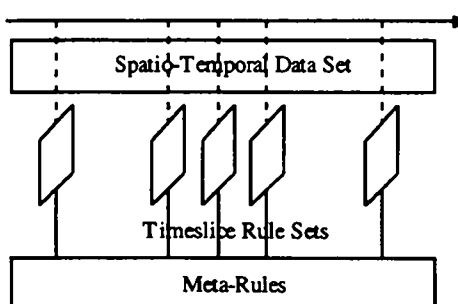

Figure 2: Spatio-Temporal Meta-rule Mining using Snapshots

The typical meta-rule extraction process consists of four steps (Abraham & Roddick 1997).

_ Select and mine individual snapshots and collect the associated rules into separate rule sets, which are also timestamped for identification purposes.

_ "Normalise", i.e., if necessary, convert the contents of the snapshot rule sets into a consistent format to facilitate further examination. This may mean, for example, that the components of rules are put into a tabular format. A general rule reading *If A then B with probability p* can thus become a three-column table with separate columns for A, B and p.

_ Compare the contents of two rule sets and separate the results into four groups: *new, expired, unchanged* and *changed* rules. To determine which category a particular rule belongs to, we use user-defined *difference*

measures. These can tell us that a rule in the later rule set belongs to the new category because it did not exist in the earlier set, or that another rule in the earlier set belongs to the expired category because it no longer exists in the later set, and so on. For example, in the above *If A the B with p* expression, we say that the combination of A and B defines a particular rule and p is selected as the difference measure. Thus, if a particular rule *If A₁ then B₁ with p₁* exists in the first rule set and changes to *If A₁ then B₁ with p₂* in the second, then it will belong to the changed category. Similarly, if the rule *If A₂ then B₂ with p₃* exists in the earlier rule set but not in the later one in its original or modified form, the rule will be placed in the expired category.

_ Conclude the meta-rule mining process by producing general descriptions for the contents of each of the four (or possibly some selected, e.g. *new* and *expired*) categories. This step constitutes the second level of abstraction described by the term meta-rule.

(b) Evolution rules

Spatio-temporal evolution rules are derived from spatio-temporal evolution descriptions that represent changes and movement of spatial objects and phenomena directly. An evolution description relates to a specific entity or group of entities, e.g. a forest fire or an animal species, and describes their spatial and attribute evolution over time. Because the information being described is not generally available explicitly in temporal SIS, the first step in producing evolution rules is the extraction of these primitive evolution descriptions from the data. This is in contrast to meta-rule mining, where pre-compiled rule sets are compared. Investigating the relationships between evolution descriptions can reveal, for example, cause and effect relationships and associations. For example, the spread of a fire can be associated with changes in wind direction, while animal movement may be affected by temperature changes or declining or increasing levels of rainfall in habitat areas or changes in sea levels. Thus, initial movement and change descriptions can be processed into rules to produce meaningful conclusions. The derivation of these rules, however, is a more complex problem than meta-rule mining and a number of sub-problems must be further investigated. Among them, the selection of predicates used in the initial evolution descriptions and the specification of techniques used to produce the final rules are some of the important ones and are currently being examined.

(c) Spatio-temporal generalisation

The attribute oriented induction technique of (Cai, et al. 1991) is a form of generalisation using conceptual hierarchies as background knowledge in the discovery process (see Figure 1 for some example hierarchies, including a spatial one). The induction technique provides a general characterisation of a selected group of items based on the commonness of their attribute values, e.g. the description *Middle-aged male professors earn more than $50,000 a year* can be produced from a university staff database. Contrasting attribute values for similar item groups in separate collections, on the other hand, can be used to derive discriminatory information that can be used for comparisons, e.g. the salaries of professors in two separate university departments.

Lu, et al. (1993) provides a spatial extension to the induction technique and further develops the generalisation concept by allowing the acquisition of characteristic information in two different ways. Both techniques are suited to analyse GIS data.

_ In spatial-data-dominant generalisation, first spatial regions are obtained by ascending a spatial conceptual hierarchy, then attribute information is generalised for each region. This may give us general descriptions of the temperature in certain regions of a map area, e.g. *The temperature in South Australia is mild.*

_ In nonspatial-data-dominant generalisation, first attribute data is generalised for all individual areas surveyed, then adjacent regions with the same characteristics are merged, yielding an overall picture of the spatial distribution of generalised attribute values. Thus, for example, if several small counties in a state have a mild temperature, they could form a 'supercounty' within the state that is joined together solely because of the shared temperature characteristic.

Both of these approaches can be used and are considered to yield *snapshot characteristics* in a temporal SIS. For mining characteristic rules that incorporate time, we have two options, which depend on the model employed by the system.

_ If snapshots are extracted for different times as described above, meta-rules can be used to describe changes between snapshot rule sets as in Figure 2.

    * In the spatial-data-dominant case, this means trends in attribute value changes can be observed, although sometimes these changes may not be detected if there are only small variations in the attribute values over time. This happens because the generalisation process can 'smooth out' deviations between layers by using high level concepts for descriptions. For example, a minor increase in temperatures over time may still be classified as *mild* in both the original and new

description.

* In the nonspatial-data-dominant case, meta-rules can describe changes in spatial areas, such as the spread of a fire or disease. In our view, this is a useful approach in many circumstances, because it enables the detection of environmental changes such as deforestation, increases or decreases in the levels of precipitation, or the spread of urban areas. As an example, consider an orchard sectioned according to a spatial hierarchy. If the occurrences of a certain tree disease are registered regularly over time, the affected areas for each survey can be found and meta-rules can be constructed to observe changes in those areas.

_ The second approach is applicable only to representational techniques where full histories of individual objects are preserved. This means that temporal intervals denote the existence of object states. To explore such data, the use of temporal generalisation is an option. At present, however, very little information is available on how to effectively generalise temporal intervals (Rainsford & Roddick 1997), and thus only hypothetical discovery guidelines can be given here.

* In the spatial-data-dominant case, for each region generated, temporal generalisation followed by attribute generalisation can be performed. This produces intervals for each region with an associated attribute concept that describes prevailing conditions for the region in a given period.

* Similarly, in the nonspatial-data-dominant case, temporal generalisation followed by attribute generalisation can precede the region merging process. This yields intervals for each of which attribute data is generalised and regions with similar characteristics are joined.

(d) Clustering

Clustering techniques used in spatial databases organise data into groups based on the similarity of their spatial features, without requiring any *a priori* knowledge, by mostly employing a distance based similarity measure. They produce a number of structures (clusters) comprising points or regions that represent spatial entities. Once these are attained, another learning algorithm can be applied to establish, for example, characteristic details of these clusters. Alternatively, attribute values of spatial objects can be generalised first and associated spatial clusters described by the same general attribute concepts may be found (Ng & Han 1994).

Time is another attribute in a multi-dimensional data space, possessing magnitude, and can be easily clustered on. The information attained by temporal clustering can be useful in e.g. detecting that certain spatial rules apply during different temporal intervals. For example, increases in a population can lead to an increase in the space taken by the population but only during spring. Spatial clustering and attribute generalisation can thus be combined with temporal clustering, but will produce different results depending on the order in which they are applied to the data. Below we give some examples (Figure 3 illustrates ways the three techniques may be combined[1]).

_ In a cadastral database, spatial clusters of a certain building type may be produced, followed by clustering on a temporal attribute such as ownership. The resulting information relates to periods of frequent sales of the pre-selected building type.

_ If we are interested in sales patterns in general in the same cadastral database, performing a temporal clustering operation first followed by either spatial clustering or attribute generalisation would produce periods of increased sales activity whilst pinpointing the areas and/or building types involved.

_ A further possibility is to use temporal clustering in conjunction with a "hashing function" that divides data into groups based on a selected criteria, e.g. day of month or time of day, enabling clusters to be found in derived (or a function of) data. For example, finding temporal clusters in the times of day for Automatic Teller Machines can give us congestion periods, which can be further processed either to find whether withdrawals, deposits or transfers dominate at these times, or to determine the spatial distribution of the machines involved.

Clusters of spatial objects can also be used in conjunction with other temporal data mining techniques, such as temporal generalisation and meta-rules. Some spatial objects, such as certain buildings, volcanos, etc., can be viewed to be spatially fixed over a long period of time. If clusters of these objects are established, it becomes possible to temporally generalise some of their attributes to characterise their behaviour over time. For other objects, the position of some spatial clusters may slowly change over time in a fashion that the identification of each cluster remains possible, such as forested areas. In this case, meta-rules may be found to describe their movement or change in size.

---

1 Temporal clustering is not necessarily limited to spatio-temporal information. Applying it to transaction databases, e.g. department store sales or bank ATM (Automatic Teller Machine) records could detect patterns of purchases/usage useful in planning and management.

Original Data Set

TC - Temporal Clustering
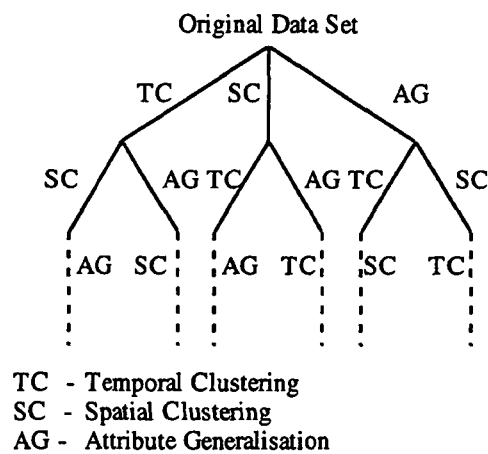SC - Spatial Clustering
AG - Attribute Generalisation

Figure 3: Spatio-Temporal Clustering

(e) Spatio-temporal associations

This special kind of probabilistic dependency information between attribute values within records is especially useful for management purposes (Agrawal, et al. 1993). Formally, it is described by the formula X Þ Y (c%, s%), meaning that the occurrence of X in a record implies the presence of Y with a confidence of c%, having at least s% of the data items supporting this claim. For example, a rule extracted from a supermarket database reading *Bread Þ Butter (90%, 15%)* suggests that 90% of purchases that include bread also include butter, with 15% of all records having recorded a combined bread and butter purchase. In practice, associations with a high level of confidence and at least a moderate amount of support are searched for. Support can be a relatively low percentage if the data set is large. The original rule extraction algorithm designed for the analysis of transaction records in (Agrawal, et al. 1993) has since been both improved and extended to different domains, with qualitative, quantitative, negative, cyclic, fuzzy, multi-level and generalised rules being searched for (Agrawal & Srikant 1994, Han & Fu 1995, Koperski & Han 1995, Ozden, et al. 1998, Savasare, et al. 1998, Srikant & Agrawal 1995, Srikant & Agrawal).

The spatial association extraction algorithm discussed by Koperski & Han (1995) searches for generalised associations between spatial objects, or spatial objects and attributes. Rules are expressed by spatial and nonspatial predicates, e.g. *schools are* close_to *parks*. Thus, they are an improvement over traditional associations with predicates now used in the antecedent and consequence. Typically, conceptual hierarchies are employed in the generalisation of spatial objects and predicates. This enables the finding of multi-level associations as both the above example and one of its possible lower level versions, *primary_schools are* close_to *parks_with_playgrounds* can be a strong rule. Similar to this spatial approach, we can think of purely temporal associations as rules extracted from temporal data and expressed by the use of at least one temporal predicate, e.g. *car oil change* precedes *tune-up*. Thus, a spatio-temporal association rule would comprise both spatial and temporal predicates. A possible method to derive these rules is by processing evolution descriptions that may already contain such predicates, as we discussed before. This enables us to find patterns such as the convergence of events in a spatial location or the seemingly unrelated movements of spatial objects caused by some common factor. Some possible examples include observing large numbers of seals near the Antarctic during summer or the movements of fishing ships towards a particular reef after coral spawning.

CONCLUSIONS

The progress in compiling spatio-temporal data sets in Spatio-Temporal Information Systems provides new opportunities in the analysis and prediction of trends in geo-referenced information. Data mining offers techniques for this purpose. In this paper, we examined the extension of established spatial discovery techniques in the spatio-temporal domain and discussed new ones that exploit the temporal aspects of this data.

Consequently, if fully developed and integrated into SIS, they offer tools for more detailed analysis capable of detecting previously unknown regularities in data. Meta-rules combined with spatial data mining techniques are suited for system implementations storing layered temporal information, such as GIS with survey data or a satellite image database. Evolution rules can be utilised in systems with more sophisticated temporal models that are continually updated when changes occur in the environment. The development of these tools, however, are

dictated by the querying requirements of the users of Spatio-Temporal Information Systems. The identification of these requirements thus remains important in helping to accommodate data mining techniques within STIS. Our current research concentrates on the development of generic knowledge discovery tools to be used on spatio-temporal data. By incorporating these tools into various types of STIS, we are aiming to provide users with enhanced analytical and data management capabilities. To this effect, we shall define a framework for mining evolution and meta-rules, and test their use in a prototype spatio-temporal knowledge discovery system.

REFERENCES

Abel, D. J. (1996) "What's Special about Spatial?", **Proceedings of 7th Australian Database Conference,** Melbourne, Australia, January 29-30.

Abraham, T. & Roddick, J. F. (1997) "Discovering Meta-Rules in Mining Temporal and Spatio-Temporal Data", **Proceedings of the 8th International Database Workshop**, Hong Kong, July 29-31.

Abraham, T. & Roddick, J. F. (1996) **Survey of Spatio-Temporal Databases**, Technical Report CIS-96-011, University of South Australia, Adelaide.

Agrawal, R., Imielinski, T. & Swami, A. (1993) "Mining Association Rules between Sets of Items in Large Databases", **Proceedings of the 1993 International Conference on Management of Data (SIGMOD 93)**, Washington, D.C., May 26-28.

Agrawal, R. & Srikant, R. (1994) "Fast Algorithms for Mining Association Rules", **Proceedings of VLDB 1994**, Santiago, Chile, September 12-15.

Al-Taha, K. K. & Barrera, R. (1990) "Temporal Data and GIS: An Overview", **Proceedings of GIS/LIS 90**, Anaheim, California, November 7-10.

Aref, W. & Samet, H. (1991) "Extending a DBMS with Spatial Operations", **Proceedings of 2nd Symposium on Spatial Database Systems**, Zürich, Switzerland, August 28-30.

Armstrong, M. P. (1988) "Temporality in Spatial Databases", **Proceedings of GIS/LIS 88**, San Antonio, Texas, November 1988.

Basoglu, U. & Morrison, J. L. (1978) "An Efficient Hierarchical Data Structure for the U.S. Historical County Boundary Data File", **Harvard Papers on GIS**, Vol 4, pp 1-21.

Bell, D. A., Anand, S. S. & Shapcott, C. M. (1994) "Database Mining in Spatial Databases", **Proceedings of the International Workshop on Spatio-Temporal Databases.**

Beller, A. (1991) "Spatial/Temporal Events in a GIS", **Proceedings of GIS/LIS 91**, Atlanta, Georgia, October 28-November 1.

Bonfatti, F. & Monari, P. D. (1993) "Spatio-Temporal Modeling of Complex Geographical Structures", **Proceedings of the IFIP TC5/WG5.11 Working Conference on Computer Support for Environmental Assessment (CSEIA 93)**, Como, Italy, October 6-8.

Cai, Y., Cercone, N. & Han, J. (1991) "Attribute-Oriented Induction in Relational Databases", G. Piatetsky-Shapiro & W. J. Frawley (eds.) **Knowledge Discovery in Databases**, AAAI Press/MIT Press.

Chen, M.-S., Han, J. & Yu, P. S. (1996) "Data Mining: An Overview from Database Perspective", **IEEE Transactions on Knowledge and Data Engineering**, December, Vol 8, pp 866-883.

Cheung, D. W., Ng, V. T. & Tam, B. W. (1996) "Maintenance of Discovered Knowledge: A Case in Multi-Level Association Rules", **Proceedings of the Second International Conference on Knowledge Discovery & Data Mining**, Portland, Oregon, August 2-4.

Easterfield, M., Newell, R. G. & Theriault, D. (1991) "Modelling Spatial and Temporal Information", **Proceedings of EGIS 91**, Brussels, Belgium, April 2.

Ester, M., Kriegel, H.-P. & Xu, X. (1995) "Knowledge Discovery in Large Spatial Databases: Focusing Techniques for Efficient Class Identification", **Proceedings of the 4th International Symposium on Large Spatial Databases (SSD 95)**, Portland, Maine, August 6-9.

Fayyad, U., Haussler, D. & Stolorz, P. (1996a) "KDD for Science Data Analysis: Issues and Examples", **Proceedings of the Second International Conference on Knowledge Discovery & Data Mining**, Portland, Oregon, August 2-4.

Fayyad, U., Piatetsky-Shapiro, G. & Smyth, P. (1996b) "Knowledge Discovery and Data Mining: Towards a Unifying Framework", **Proceedings of the Second International Conference on Knowledge Discovery & Data Mining**, Portland, Oregon, August 2-4.

Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P. & Uthurusamy, R. (1996c) **Advances in Knowledge Discovery and Data Mining**, AAAI Press/MIT Press, Menlo Park.

Guenther, O. & Buchmann, A. (1990) "Research Issues in Spatial Databases", **SIGMOD Record**, December, Vol 19, pp 61-68.

Güting, R. H. (1994) "An Introduction to Spatial Database Systems", **VLDB Journal**, October, Vol 3, pp 357-399.

Hachem, N. I., Gennert, M. A. & Ward, M. O. (1992) "A DBMS Architecture for Global Change Research", **Proceedings of ISY Conference on Earth and Space Science**, Pasadena, California, February.

Han, J. (1996) "Data Mining Techniques", **Proceedings of ACM-SIGMOD 96 Conference**, Montreal, Canada, June 4-6.

Han, J. & Fu, Y. (1995) "Discovery of Multiple-Level Association Rules from Large Databases", **Proceedings of VLDB 1995**, Zürich, Switzerland, September 11-15.

Han, J., Koperski, K. & Stefanovic, N. (1997) "GeoMiner: A System Prototype for Spatial Data Mining", **Proceedings of the 1997 ACM-SIGMOD Conference on the Management of Data**, Tucson, Arizona, May 12-15.

Hazelton, N. W. J. (1992) "Beyond the 2-D Map: A New Metaphor for Multi-Temporal 4-D GIS.", **Proceedings of GIS/LIS 92**, San Jose, California, November 10-12.

Holsheimer, M. & Siebes, A. P. J. M. (1994) **Data Mining: The Search for Knowledge in Databases**, Technical Report CS-R9406, CWI, The Netherlands.

Kemp, Z. & Kowalczyk, A. (1994) "Incorporating the Temporal Dimension in a GIS" M. F. Worboys (eds.) **Innovations in GIS**, Taylor & Francis.

Koperski, K., Adhikary, J. & Han, J. (1996) "Spatial Data Mining: Progress and Challenges Survey Paper", **Proceedings of 1996 ACM-SIGMOD Workshop on Research Issues on Data Mining and Knowledge Discovery**, Montreal, Canada, June 1996.

Koperski, K. & Han, J. (1995) "Discovery of Spatial Association Rules in Geographic Information Databases", **Proceedings of the 4th International Symposium on Large Spatial Databases (SSD 95)**, Portland, Maine, August 6-9.

Langran, G. (1992) **Time in Geographic Information Systems**, Taylor & Francis, London.

Langran, G. & Chrisman, N. R. (1988) "A Framework for Temporal Geographic Information", **Cartographica**, Vol 25, pp 1-14.

Lu, W., Han, J. & Ooi, B. C. (1993) "Discovery of General Knowledge in Large Spatial Databases", **Proceedings of the 1993 Far East Workshop on GIS (IEGIS 93)**, Singapore, June 21-22.

Mannila, H. (1997) "Methods and Problems in Data Mining", **Proceedings of the International Conference on Database Theory**, Delphi, Greece, January 8-10.

Medeiros, C. B. & Jomier, G. (1994) "Using Versions in GIS", **Proceedings of DEXA 94**, September 7-9.

Medeiros, C. B. & Pires, F. (1994) "Databases for GIS", **SIGMOD Record**, March, Vol 23, pp 107-115.

Muntz, R., Shek, E. & Zaniolo, C. (1995) "Using LDL++ for Spatio-Temporal Reasoning in Atmospheric Science Databases" R. Ramakrishnan (eds.) **Applications of Logic Databases**, Kluwer Academic.

Neteler, M. (1997) **Introduction to GRASS GIS Software**, Institute of Physical Geography and Landscape Ecology, University of Hannover, Germany.

Ng, R. T. (1996) "Spatial Data Mining: Discovering Knowledge of Clusters from Maps", **Proceedings of 1996 ACM-SIGMOD Workshop on Research Issues on Data Mining and Knowledge Discovery**, Montreal, Canada, June 1996.

Ng, R. T. & Han, J. (1994) "Efficient and Effective Clustering Methods for Spatial Data Mining", **Proceedings of VLDB 1994**, Santiago, Chile, September 12-15.

Ozden, B., Ramaswamy, S. & Silberschatz, A. (1998) "Cyclic Association Rules", **Proceedings of the 14th International Conference on Data Engineering (ICDE `98)**, Orlando, Florida, February 23-27.

Paredaens, J. (1995) "Spatial Databases, The Final Frontier", G. Gottlob & M. Y. Vardi (eds.) **Database Theory**, Springer-Verlag.

Peuquet, D. & Wentz, E. (1994) "An Approach for Time-Based Analysis of Spatiotemporal Data", **Proceedings of the 6th International Symposium on Spatial Data Handling**, Edinburgh, Scotland, September 5-9.

Peuquet, D. J. & Duan, N. (1995) "An Event-based Spatiotemporal Data Model (ESTDM) for Temporal Analysis of Geographical Data", **International Journal of Geographical Information Systems** January, Vol 9, pp 7-24.

Piatetsky-Shapiro, G. & Frawley, W. J. (1991) **Knowledge Discovery in Databases**, AAAI Press/MIT Press, Menlo Park.

Quinlan, J. R. (1986) "The Effect of Noise on Concept Learning", Michalski, R. S., Carbonell, J. G. & Mitchell, T. M. (eds.), **Machine Learning, an Artificial Intelligence Approach, Vol 2**, Morgan Kaufmann, San Mateo, California.

Rafaat, H., Yang, Z. & Gauthier, D. (1994) "Relational Spatial Topologies for Historical Geographical Information", **International Journal of Geographical Information Systems**, March, Vol 8, pp 163-173.

Rainsford, C. P. & Roddick, J. F. (1997) "The Attribute Oriented Induction of Rules from Temporal Interval Data", **Proceedings of the 8**[th] **International Database Workshop**, Hong Kong, July 29-31.

Rainsford, C. P. & Roddick, J. F. (1996) **A Survey of Issues in Data Mining**, Technical Report CIS-96-006, University of South Australia, Adelaide.

Raper, J. (1995) "Making GIS Multidimensional", **Proceedings of the Joint European Conference and Exhibition on Geographical Information (JEG-CI 95)**, The Hague, The Netherlands, March 26-31.

Rojas-Vega, E. & Kemp, Z. (1995) "An Object-Oriented Data Model for Spatio-Temporal Data", **Proceedings of the Ninth Annual Symposium on Geographic Information Systems**, Vancouver, Canada, March 27-30.

Roshannejad, A. A. & Kainz, W. (1995) "Handling Identities in Spatio-Temporal Databases", **Proceedings of ACSM/ASPRS 1995 Annual Convention and Exposition Technical Papers**, Charlotte, North Carolina,

Savasare, A., Omiecinski, E. & Navathe, S. (1998) "Mining for Strong Negative Associations in a Large Database of Customer Transactions", **Proceedings of the 14**[th] **International Conference on Data Engineering (ICDE '98)**, Orlando, Florida, February 23-27.

Sinton, D. F. (1978) "The Inherent Structure of Information as a Constraint to Analysis: Mapped Thematic Data as a Case Study", **Harvard Papers on GIS**, Vol 7, pp 1-17.

Srikant, R. & Agrawal, R. (1995) "Mining Generalized Association Rules", **Proceedings of VLDB 1995**, Zürich, Switzerland, September 11-15.

Srikant, R. & Agrawal, R. (1996) "Mining Quantitative Association Rules in Large Relational Tables", **Proceedings of the 1996 ACM-SIGMOD Conference on the Management of Data**, Montreal, Canada, June 4-6.

Stolorz, P., Nakamura, H., Mesrobian, E., Muntz, R. R., Shek, E. C., Santos, J. R., Yi, J., Ng, K., Chien, S.-Y., Mechoso, C. R. & Farrara, J. D. (1995) "Fast Spatio-Temporal Data Mining of Large Geophysical Sets", **Proceedings of the 1st International Conference on Knowledge Discovery and Data Mining**, Montreal, Canada, August 20-21.

Story, P. A. & Worboys, M. F. (1995) "A Design Support Environment for Spatio-Temporal Database Applications", A. U. Frank & W. Kuhn (eds.) **Spatial Information Theory: A Theoretical Basis for GIS**, Springer-Verlag.

Wachowicz, M. & Healey, R. G. (1994) "Towards Temporality in GIS", M. F. Worboys (eds.) **Innovations in GIS**, Taylor & Francis.

Wang, W., Yang, J. & Muntz, R. (1997) "STING: A Statistical Information Grid Approach to Spatial Data Mining", **Proceedings of VLDB 97**, Athens, Greece, August 26-29.

Worboys, M. F. (1992) "Object-Oriented Models of Spatiotemporal Information", **Proceedings of GIS/LIS 92**, San Jose, California, November 10-12.

Yeh, T. S. & Viémont, Y. H. (1992) "Temporal Aspects of Geographical Databases", **Proceedings of EGIS 92**, Munich, Germany, March 23-26.

Yuan, M. (1996) "Incorporating Human Conceptualizations into GIS Representation to Support Spatiotemporal Queries", Personal Communication, November.

Zhang, T., Ramakrishnan, R. & Livny, M. (1996) "BIRCH: An Efficient Clustering Method for Very Large Databases", **Proceedings of 1996 ACM-SIGMOD Workshop on Research Issues on Data Mining and Knowledge Discovery**, Montreal, Canada, June 1996.