Spatial, Temporal and Spatio-Temporal Databases - Hot Issues and Directions for PhD Research

John F. Roddick
School of Informatics and Engineering
Flinders University
GPO Box 2100, Adelaide 5001
South Australia
roddick@infoeng.flinders.edu.au

Max J. Egenhofer
National Center for Geographic Information and Analysis
University of Maine
Orono, Maine 04469-5711
USA
max@spatial.maine.edu

Erik Hoel
Research and Development
Environmental Systems Research Institute
380 New York Street
Redlands, California 92373 USA
ehoel@esri.com

Dimitris Papadias
Department of Computer Science
Hong Kong University of Science and Technology
Clearwater Bay
Hong Kong
dimitris@cs.ust.hk

Betty Salzberg
College of Computer and Information Science
Northeastern University
Boston, Massachusetts 02115
USA
salzberg@ccs.neu.edu

Abstract

Spatial and temporal database systems, both in theory and in practice, have developed dramatically over the past two decades to the point where usable commercial systems, underpinned by a robust theoretical foundation, are now starting to appear. While much remains to be done, topics for research must be chosen carefully to avoid embarking on impractical or unprofitable areas. This is particularly true for doctoral research where the candidate must build a tangible contribution in a relatively short time.

The panel session at the Eighth International Symposium on Spatial and Temporal Databases (SSTD 2003) held on Santorini Island, Greece [7] in July 2003 thus took as its focus the question *What to focus on (and what to avoid) in Spatial and Temporal Databases : recommendations for doctoral research*. This short paper, authored by the panel members, summarizes these discussions.

1 Introduction

The field of spatial and temporal database systems has captured the imagination of many researchers including many current and potential doctoral students. It represents the promise of being able to contribute in both theoretical and practical terms to an area of interest to industry and with sufficient scope for a reasonable period of academic investigation.

Nevertheless, picking (or recommending) a doctoral topic can be difficult. Candidates often run the risk of embarking on research that is either too trivial, in terms of theoretical contribution, or too large in terms of the practical implementation that is required to validate an idea. In particular, at this stage in the development of spatial and temporal database systems research there are dangers that candidates could pick a topic that:

- represents a career rather than a thesis by failing to specialize sufficiently. While the field is relatively young, substantial research already exists and the area is sufficiently complex to require such specialization;
- consists primarily of an implementation of other re-

searchers' good ideas—a particular risk given the industry-pull for artifacts in spatial and temporal database systems; and

- is a transiently popular area which, by the end of the thesis, will not result in a contribution worthy of a doctorate. The area is moving quickly and *bandwagon* topics are quickly solved.

In common with many areas, we believe that spatial and temporal database systems research should generally be characterized by research that:

- is collaborative with other research in cognate areas of information technology research, such as mobile telecommunications;
- integrates with other disciplines such as medicine, CAD /CAM, GIS, environmental science, molecular biology or genomics/bioinformatics;
- uses large, real databases—synthetic databases are often of less use in this area than in many others;
- is interesting—both to maintain a candidate's enthusiasm through the three or more years of candidature and which is of practical interest to industry and/or society; and
- where possible, attempts to change the accepted paradigm or approach to a given problem.

Interestingly, there was a general consensus that spatial and temporal database doctoral research benefits from aiming at the production of high quality papers rather than at the production of a thesis *per se*. Very often a Ph.D. thesis is compiled from related papers the student has previously written and submitted to conferences and journals.

A long discussion took place on how much a doctoral student should read. It was felt that students should learn the basic techniques and read *some* of the more important background papers and a small number of papers directly relevant to the student's work. However, it was suggested that a student can get distracted and depressed and not make progress if paper-reading goes on too long. Sometimes it is better to write something and find out later that similar work has been done rather than to write nothing. Thesis advisors have an important role in guiding the student to appropriate problems to get started on.

2 Specific Technological Recommendations

As well as the general principles outlined in Section 1, a number of specific recommendations were made regarding topics to consider and topics to avoid. These are discussed below and are in no particular order.

2.1 Topics to Consider

Information extraction from spatial data sources.
Useful spatial and temporal database systems applications are predicated on access to reliable and up-to-date source data. However, while disk storage availability, CPU power and techniques for interpreting the semantic content of many forms of data have advanced, techniques for extracting spatial information have lagged behind. Moreover, the modelling of spatial data sources is as important as the data themselves if such data are to be interpreted correctly and the development of appropriate models for spatial data streams is an open problem. This lack of data and data model are together rendering many good ideas and prototype systems impractical.

Models that survive the real-world.
The implementation of systems in the real-world is difficult and many systems have proven impractical for reasons such as:

- minor but significant discrepancies between the real-world and the modelled environment (for example, road closures or changes in museum opening times);
- the loss of mobile communication coverage for systems that rely on access to central servers;
- a reliance on particular system configurations or, more importantly, a particular pattern of behaviour of the user;
- an inability to self-optimise for different temporal cycles and significant changes in spatial density;
- trade-offs between complexity and performance for real-world models.

Terrain models.
Terrain can be considered the *step-child* of spatial and temporal database systems research. Many existing spatial models do not translate well to terrain and there is a feeling that a fresh start is needed [5]. Of particular importance here is to reconceptualize the ontological properties of terrain and the operations that should be available over such models.

'Natural'/usable user interfaces and the visualization of spatio-temporal data.
Implementation of many spatial and temporal database systems to date has demonstrated the limitations of existing user interfaces in real-world situations. The standard WIMP interface has significant drawbacks and ideas such as augmented reality systems with head-up displays and magic wands [2] that can be used to point at real-world objects would be useful devices.

Coupled with this is the need for research into effective techniques for visualizing spatio-temporal data in the context of both static and animated graphics/maps.

Spatio-temporal data mining.
Data mining and knowledge discovery have become popular fields of research. A significant subset of this research is looking at the particular semantics of space and time and the manner in which they can be sensibly accommodated into data mining algorithms [12]. Most of the work, although certainly not all, can be placed in one of three catagories:

- Temporal Association Rule Mining, which aims to detect correlations in transactional and relational data that possess a time component. Some (more limited) work has investigated Spatial Association Rule Mining.
- Spatial Clustering, which aims to group similar objects into the same cluster while grouping dissimilar objects into different clusters. In this case similarity is influenced by both spatial and aspatial attributes of the objects as well as any obstacles that may exist.
- Time-series Analysis, which aims to detect frequent patterns in the values of an attribute over time.

Significantly, most of this work deals with one or the other of spatial or temporal semantics, with very few handling both.

Spatio-temporal applications for mobile, wireless, location-aware services and sensor networks.
Two technologies that have emerged recently offer substantial scope for application development – wireless-based, location-aware devices and networks of sensors [10].

While researchers should avoid particular technologies in this area (such as relying on a particular network protocol) there is substantial scope for systems that leverage the mobility of the user in a way not possible before. Such systems will almost inevitably require spatial database support. Sensor networks and streaming data are very hot topics right now and the field is rapidly developing. It will take an effort to keep up with developments in this area.

Spatio-temporal modelling as a network of cells.
The recent papers [1] and [6] use fixed-boundary cells to model spatio-temporal data. This may be a 'back-to-the-future' solution as certainly the use of fixed grids predated R-trees. R-trees are often chosen for indexing spatial and spatio-temporal data because the code is available and the method is well-understood. However, the overlapping nature of R-trees, which causes backtracking in search, may not be best suited for every application. There may be many situations where non-overlapping search structures are both more natural and more efficient.

Spatio-temporal vacuuming.
While disk storage cost is decreasing and data storage capacity is increasing, there is still a need to delete obsolete data. Vacuuming is used to delete data that is no longer of interest and has been investigated in terms of temporal databases [9]. Extending this to the development of models of obsolescence and to cater for spatial and spatio-temporal data is still required.

Unconventional spatio-temporal access methods.
Unconventional problems call for fresh approaches. As an example, much of the work on moving objects in networks (see for example [11] and [8]) uses graphs to model the spatial area instead of using Euclidean space representations. Road distances may be used rather than Euclidean distances. The access method must match the problem. Research on spatio-temporal access methods has mainly focused on two aspects: (i) storage and retrieval of historical information, and (ii) future prediction. Several indexes, usually based on multi-version or 3-dimensional variations of R-trees, have been proposed towards the first goal, aiming at minimizing storage requirements and query cost. Methods for future prediction assume that, in addition to the current positions, the velocities of moving objects are known. The goal is to retrieve the objects that satisfy a spatial condition at a future timestamp (or interval) given their present motion vectors (e.g., 'based on the current information, find the cars that will be in the city center 10 minutes from now'). The only practical index in this category is the TPR-tree and its variations (also based on R-trees).

Despite the large number of methods that focus explicitly on historical information retrieval or future prediction, currently there does not exist a single index that can achieve both goals. Even if such a 'universal' structure existed (e.g., a multi-version TPR-tree keeping all previous history of each object), it would be inapplicable for several update-intensive applications, where it is simply infeasible to continuously update the index and at the same time process queries. For instance, an update (i.e., deletion and re-insertion) in a TPR-tree may need to access more than 100 nodes, which means that by the time it terminates its result may already be outdated (due to another update of the same object).

Even for a small number of moving objects and a low update rate, the TPR-tree (or any other index) cannot 'follow' the fast changes of the underlying data. Consequently, main-memory structures seem more appropriate for update-intensive applications.

Novel query types and space/time queries.
Jim Gray gave a SIGMOD Record interview [14] that provided a context for thinking about 'management of data' beyond the context of SQL. For database researchers, there is more to life than SQL—there are very interesting problems that can only be solved with loosely-coupled system architectures (i.e., web services, grid computing) and that there are other kinds of transactions beyond the classic Codd short transaction. Note that this interview also provides insight into finding interesting research topics for fledgling researchers.

An example of novel query type in spatio-temporal databases refers to continuous queries whose result is strongly related to the temporal context. An example of a continuous spatio-temporal query is: 'based on my current direction and speed of travel, which will be my nearest two gas stations for the next 5 minutes?'. A result of the form $<\{A, B\}, [0, 1)>$, $<\{B, C\}, [1, 5)>$ would imply that A, B will be the two nearest neighbors during interval $[0, 1)$, and B, C afterwards. Notice that the corresponding instantaneous query ('which are my nearest gas stations now?') is usually meaningless in highly dynamic environments; if the query point or the database objects move, the result may be invalidated immediately. Any spatial query has a continuous counterpart whose termination clause depends on the user or application needs. Consider, for instance, a window query, where the window (and possibly the database objects) moves/changes with time. The termination clause may be temporal (for the next 5 minutes), a condition on the result (e.g., until exactly one object appears in the query window, or until the result changes three times), a condition on the query window (until the window reaches a certain point in space) etc. A major difference from continuous queries in the context of traditional databases is that, in case of spatio-temporal databases, the object's dynamic behavior does not necessarily require updates, but can be stored as a function of time using appropriate indexes. Furthermore, even if the objects are static, the results may change due to the dynamic nature of the query itself (i.e., moving query window), which can be also represented as a function of time. Thus, a spatio-temporal continuous query can be evaluated instantly (i.e., at the current time) using time-parameterized information about the dynamic behavior of the query and the database objects, in order to produce several results, each covering a validity period in the future [13].

Moving object tracking research.
Data logging and visualization is a simple problem, however, the development of analytical techniques for integrating moving and static datasets is necessary. One example can be found in [13].

Approximate queries.
In several spatio-temporal applications, the size of the data and the rapidity of updates necessitate approximate query processing. For instance, in traffic supervision systems, the incoming data are usually in the form of data streams (e.g., through sensors embedded on the road network), which are potentially unbounded in size. Therefore, materializing all data is unrealistic. Furthermore, even if all the data were stored, the size of the index would render exact query processing very expensive since any algorithm would have to access at least a complete path from the root to the leaf level. Finally, in several applications the main focus of query processing is retrieval of approximate summarized information about objects that satisfy some spatio-temporal predicate (e.g., 'the number of cars in the city center 10 minutes from now'), as opposed to exact information about the qualifying objects (i.e., the car ids), which may be unavailable, or irrelevant.

Parallel spatio-temporal algorithms.
With the overwhelming volumes of spatio-temporal data available in the real-world, the importance of parallel algorithms that utilize coarse-grained parallel processing environments remains undiminished. However, Jim Gray has noted in [14] that because of a number of changes in hardware and software and caching algorithms, bigger and bigger computers have been built. So one must be cautious in assuming that today's larger machines cannot solve the problem.

Uncertainty.
Uncertainty is inherent is most spatio-temporal applications due to measurement/ digitization errors and missing or incomplete information. Assume, for instance, a user with a PDA inquiring about the closest restaurant in terms of road network distance. Although the user may actually be on a road segment, due to the inaccuracy of the GPS device, the system may fail to recognize this. Such a situation can be handled by defining an (application-dependent) threshold d_T, so that if a point is within distance d_T from a road segment, it is assumed to lie on it. Alternatively, we can snap the point to the closest road assuming incomplete information (e.g., an un-recorded alley), or we can consider it unreachable depending on the application specifications. Similar problems exist for object trajec-

tories because while movement is continuous, measurements are discrete.

Creating and managing complex spatio-temporal simulation models.
Complex simulation modes with multiple inputs/parameters and multiple outputs remains a significant challenge. Common examples of such a model are the landuse/transportation model of a city and predictive models of environmental and social change.

Ontologies for spatio-temporal modelling.
The majority of past research in the area of spatial and spatio-temporal modeling has focused almost exclusively on geometric aspects. There is a plethora of models that deal with partitions, hierarchies, topological and direction relations, and trajectories of moving points. When modeling real-world phenomena, these geometric abstractions capture only one part of the spatial information typically used by people when making decisions. The lack of a systematic treatment of the semantics for which these geometries stand limits the analytical power, reducing the ability to integrate spatial and spatio-temporal search and retrieval methods seamlessly into the broader setting of the Web. With the push for the Semantic Web, and its spatial cousin in the form of the Geospatial Semantic Web [4], much more attention needs to be paid to spatial and spatio-temporal ontologies. The development of pilot ontologies will help, but the goal is much more to identify generic methods that support spatial search and spatio-temporal integration with the help of such ontologies.

Privacy / legal issues.
The potential for abuse in, for example, tracking individuals through their location-aware services is matched by the benefits that could be generated by such services. Outside of spatial and temporal database research, the abuse of data collected about individuals is creating alarm in some areas and the inappropriate use of personal data can cause adverse legislative reactions which may block not only illegitimate uses of such data but also potentially useful applications. The investigation of appropriate legislative frameworks is thus seen as a useful area of research.

2.2 Topics to Avoid

Technology dependent topics.
Topics that rely on the stability of a technology are doomed to obsolescence. Significantly this maybe within the period of candidature leaving a student with a well-researched but worthless thesis. Particular aspects to be careful of include research that relies on a particular platform, a transient situation in terms of a commercial situation, or performance characteristics, or even a particular standard.

New generic temporal data models.
While spatial and temporal database research is a relatively young area, some areas can be considered solved. One of these is the development of generic temporal data models. The impressive body of work by many researchers in the 1990s has laid a framework for current temporal applications including extensions to standards such as SQL. While little further work is required in this area, the consideration of particular spatial and spatio-temporal semantics is still needed.

Spatial research unable to be scaled to 3D spatial or spatio-temporal.
A particular flaw in some research is the development of techniques and tools that are simply unable to cope when expanded to handle additional dimensions. For example, 2D spatial systems unable, because of their underlying model and techniques, to handle 3D space or, more commonly, static spatial systems unable to be extended to accommodate time.

Even the R-tree has problems with 4 dimensions. Partly this is due to the fact that each index entry in an R-tree must contain the low and high boundaries in each dimension for each child. With 4 dimensions, this is 8 numbers per child. This is one reason why fixed-cell like access methods, or access methods based on kd-tree-like splits, may prove more efficient for 3D spatial or spatio-temporal problems.

'We did this' research without a generalizable lesson.
As impressive as some implementations might be, an implementation does not represent doctoral research if generalisable lessons cannot be elicited from the experience. Doctoral studies in spatial and temporal databases should rather focus on fundamental questions, where during the course of investigations prototype implementations are used as proof of concept, but certainly implementations should not be considered the end product of a dissertation.

Research that only considers time as a date or a location as a zipcode.
It is somewhat depressing to see papers claiming to be handling space and time dealing with dates as numeric values and locations in terms of discrete location identifiers such as zipcodes. Such research does not capture the semantics of space and time and thus does not represent spatial and temporal database research.

Research that is based upon absurd simplifying assumptions.

It is sometimes the case that research efforts are based upon simplifying assumptions that render the result meaningless. For example, mobile systems that require nearly infinite communication bandwidth or localized processing power. Additionally, we sometimes observe research that is based upon invasive protocols that ignores any real-world privacy concerns. While researchers are at liberty to make such assumptions, they should exercise some common sense reasoning when establishing their assumptions so as to maximize the utility of their research.

3 Final Comments

Spatial and temporal databases are an exciting and rapidly advancing field and we have outlined above a few areas we consider worthwhile for doctoral candidates just starting their research career. Please be aware, however, that the technology areas above do not represent an exhaustive list of the possible doctoral research areas in spatial and temporal databases. Moreover, our list of topics to avoid should be interpreted sensibly as there may be circumstances where useful research outcomes can be obtained.

4 Acknowledgements

The authors of the paper include those that formally presented at the panel session at the symposium (be it in person or, as in the case of Max Egenhofer, by prerecorded movie [3]). Many of the symposium delegates also contributed greatly to the debate and we would like to thank all other attendees for their contribution.

References

[1] V.P. Chakka, A.C. Everspaugh, and J.M. Patel. Indexing large trajectory data sets with SETI. *CIDR*, 2003.

[2] M.J. Egenhofer and W. Kuhn. Beyond desktop GIS. In *GIS PlaNET*, Lisbon, Portugal, 1998. `http://www.spatial.maine.edu/~max/BeyondDesktopGIS.pdf`.

[3] M.J. Egenhofer. Recommendations for Ph.D. research in spatial and temporal databases. In *Pre-recorded panel presentation at SSTD 2003.* http://www.spatial.maine.edu/~max/SSTD2003-panel-web.mov.

[4] F. Fonseca and A. Sheth. The geospatial semantic web. Research priority white paper, University Consortium for Geographic Information Science, 2003. `http://www.personal.psu.edu/faculty/f/u/fuf1/Fonseca-Sheth.pdf`.

[5] W.R. Franklin. Computational and geometric cartography. In *Keynote at GIScience 2002*, Boulder, CO, USA, 2002.`http://www.ecse.rpi.edu/Homepages/wrf/research/giscience2002/`.

[6] M. Hadjieleftheriou, G. Kollios, D. Gunopulos, and V. Tsotras. On-line discovery of dense areas in spatio-temporal databases. *In [12]*, pages 306–324.

[7] T. Hadzilacos, Y. Manolopoulos, J.F. Roddick, and Y. Theodoridis, eds. *Advances in Spatial and Temporal Databases. Proceedings of the Eighth International Symposium on Spatial and Temporal Databases, SSTD 2003, Santorini Island, Greece*, volume 2750 of *Lecture Notes in Computer Science*. Springer, Berlin, 2003.

[8] C. Jensen, J. Kolar, T. B. Pedersen, and I. Timko. Nearest neighbor queries in road networks. *GIS*, 2003.

[9] C.S. Jensen. Vacuuming. In R.T. Snodgrass, ed, *The TSQL2 Temporal Query Language*, pages Chapter 23, pp. 451–462. Kluwer Academic Publishing, Boston, 1995.

[10] S. Nittel, A. Stefanidis, I. Cruz, M. Egenhofer, D. Goldin, A. Howard, A. Labrinidis, S. Madden, A. Voidard, and M. Worboys. Report from the first workshop on geo-sensor networks. *SIGMOD Record*, 33(1), 2003.

[11] D. Papadias, J. Zhang, N. Mamoulis, and Y. Tao. Query processing in spatial network databases. *VLDB*, 2003.

[12] J.F. Roddick, K. Hornsby, and M. Spiliopoulou. YABTSSTDMR - yet another bibliography of temporal, spatial and spatio-temporal data mining research. In K.P. Unnikrishnan and R. Uthurusamy, eds, *SIGKDD Temporal Data Mining Workshop*, pages 167–175, San Francisco, CA, 2001. ACM.

[13] Y. Tao, D. Papadias, and Q. Shen. Continuous nearest neighbor search. *VLDB*, 2002.

[14] M. Winslett. Distinguished database profiles: Interview with Jim Gray. *SIGMOD Record*, 32(1), 2003.