A Bibliography of Temporal, Spatial and Spatio-Temporal Data Mining Research

John F. Roddick
School of Computer and Information Science
University of South Australia, The Levels Campus
Mawson Lakes, South Australia 5095, Australia.
+61 8 8302 3463
roddick@cis.unisa.edu.au

Myra Spiliopoulou
Institut für Wirschaftsinformatik
Humboldt-Universität, Berlin
Spandauer Str. 1, D-10178 Berlin, Germany
+49.30.2093-5730
myra@wiwi.hu-berlin.de

ABSTRACT

With the growth in the size of datasets, data mining has recently become an important research topic and is receiving substantial interest from both academia and industry. At the same time, a greater recognition of the value of temporal and spatial data has been evident and the first papers looking at the confluence of these two areas are starting to emerge. This short paper provides a few comments on this research and provides a bibliography of relevant research papers investigating temporal, spatial and spatio-temporal data mining.

Keywords

Temporal Data Mining, Temporal Rules, Temporal Patterns, Spatial Data Mining, Spatio-Temporal Data Mining.

1. INTRODUCTION

The development of data mining has naturally led to the exploration of application domains within which data mining may be used. Since many of these domains have an inherently temporal or spatial context, the time and/or space component must be taken into account in the mining process, in order to correctly interpret the collected data.

This paper presents a bibliography of those papers that have extended existing, or developed new data mining techniques to accommodate space and time. The bibliography is organized into contributions for temporal, spatial and spatio-temporal data mining.

A few words on the selection of papers in this collection are appropriate. Since data mining is an application-oriented research domain, there are many significant works that report on the results of applying a mining technique to solve a specific problem. In this collection, we focus on the analysis of spatial and temporal data as a *generic* issue, rather than on individual case studies.

Since we are reporting about research in a domain of intense activity, such a bibliography is unlikely to be complete, even at the time of submission. It is possible that we have omitted important papers, and we would be grateful for any subsequent inclusions. Our bibliography is continually updated as further relevant work is published and can be found at:

`http://www.cis.unisa.edu.au/~cisjfr/STDMPapers`

2. TEMPORAL DATA MINING

Temporal data mining concerns the analysis of events ordered by one or more dimensions of time. (Multiple dimensions of time are possible if a system accommodates multiple time-lines for an event such as valid-time, transaction-time or decision-time.) We distinguish between two broad directions. One concerns the discovery of causal relationships among temporally-oriented events. The other concerns the discovery of similar patterns within the same time sequence or among different time sequences. This latter area, commonly termed *time series analysis* (or *trend analysis*) focuses on the identification of similar *pre-specified* patterns, similar peaks for example, among time series.

Time series analysis has been a field of active research for a long time. Research aspects include curve approximation with mathematical methods; noise reduction; comparison of time series using pattern matching techniques and prediction using mathematical methods or heuristics (such as neural networks).

We address the problem of time series comparison for trend discovery in subsection 2.4. The large domain of time series description and prediction is outside the scope of our survey. A paper collection focussing on time series analysis for prediction is as follows:

Andreas S. Weigend and Neil A. Gershenfeld, editors. *Time Series Prediction: Forecasting the Future and Understanding the Past*, Volume XV of Proc. *NATO Advanced Research Workshop on Comparative Time Series Analysis*, Santa Fe, New Mexico, May 1993. Addison-Wesley.

We focus here on the discovery of causal relationships. The contributions we show are, at a first glance, fairly disparate. They encompass the discovery of temporal rules, of sequences and of patterns. However, in many respects this is just a terminological heterogeneity among researchers that are, nevertheless, addressing the same problem, albeit from different starting points and domains. Although we do not have the ambition in this paper of building a uniform terminological basis, we believe it is appropriate to bring those contributions together.

Different terminologies for the same problem.

The domain of temporal mining focuses on the discovery of causal relationships among events that may be ordered in time and may be causally related. In non-temporal data mining, this corresponds to the discovery of "rules", be they association rules, classification rules, characterisations or some other rule structure and semantics. This gives rise to terms such as "temporal rule discovery".

Ordered events form sequences and the cause of an event always occurs before its result. Thus, some of the temporal rules of interest are those that respect the order of events on the time axis. Hence, if we focus on rules that contain only conjunctions (ie. no disjunctions and no negation) we perform data mining on sequences. This gives rise to terms such as "sequence mining".

In other data mining research, we are interested in groups of items, which, taken together, satisfy certain statistical properties. The groups are comprised of events ordered by time and effort is directed towards extracting descriptions for the sequences contained in the dataset. In the terminology used in the domains relevant to such sequence processing (such as signal processing, string processing), we would say that we are looking for patterns in the sequences. This gives rise to terms like "pattern discovery".

Despite the terminological differences among the contributions, the fundamental problem being addressed by the published research using those three terms is the same.

The same terminology for different problems.

Although many contributions to this type of temporal mining use the term "pattern discovery" it should be clear that the discovery of patterns is a not peculiar to temporal mining. In software engineering, data mining techniques are used to discover similar software modules. In workflow management, similarities in process workflows are of interest. Similarly, the discovery of sequences in genome database is a further subject of intensive research.

This fact has raised the issue of which contributions should be included in this bibliography. The literature on pattern discovery in general is relevant to the discovery of causal relationships among events. However, the notion of pattern for a piece of software code is fundamentally different than for a sequence of events on the time axis. On the other hand, some of the research on pattern discovery that does focus on time sequences exploit only the most rudimentary characteristics of time, such as the total ordering of events.

Hence, we have decided to include those contributions that either (i) discover temporal rules by explicitly taking the characteristics of the temporal dimension into account or (ii) mine sequences of events on a time axis. Those requirements exclude techniques designed for pattern discovery on multi-dimensional spatial data, such as genome sequences, and in software engineering. At the same time, sequence mining techniques intended for any type of event sequences are included, regardless of the application domain (which ranges from market analysis to the mining of web traversals).

We have organized the temporal mining section of this bibliography into (1) formulation of frameworks, where general issues, including theoretical issues, relevant to both of the two broad directions are discussed, (2) analysis of causal and temporal relationships, and (3) discovery of patterns among time series. *Studies making major contributions to more than one area are repeated in each area.*

2.1 Frameworks for Temporal Mining

Al-Naemi, S. 1994. 'A Theoretical Framework for Temporal Knowledge Discovery'. In *Proc. International Workshop on Spatio-Temporal Databases*, Benicassim, Spain. 23-33.

Berger, G. and Tuzhilin, A. 1998. 'Discovering unexpected patterns in temporal data using temporal logic'. In *Temporal Databases - Research and Practice.* O. Etzion, S. Jajodia and S. Sripada (eds.), Lecture Notes in Computer Science 1399, Springer-Verlag, Berlin. 281-309.

Chen, X. and Petrounias, I. 1998. 'A framework for temporal data mining'. In *Proc. Ninth International Conference on Database and Expert Systems Applications, DEXA'98*, Vienna, Austria. Springer-Verlag, Berlin. Lecture Notes in Computer Science 1460:796-805.

Rainsford, C.P. and Roddick, J.F. 1996. 'Temporal data mining in information systems: a model'. In *Proc. Seventh Australasian Conference on Information Systems*, Hobart, Tasmania. 2:545-553.

2.2 Discovery of Causal and Temporal Rules

Abraham, T. and Roddick, J.F. 1997. 'Discovering meta-rules in mining temporal and spatio-temporal data'. In Proc. Eighth International Database Workshop, Data Mining, Data Warehousing and Client/Server Databases (IDW'97), Hong Kong. Springer-Verlag. 30-41.

Abraham, T. and Roddick, J.F. 1999. 'Incremental meta-mining from large temporal data sets'. In Advances in Database Technologies, Proc. First International Workshop on Data Warehousing and Data Mining, DWDM'98. Y. Kambayashi, D.K. Lee, E.-P. Lim, et al. (eds.), Lecture Notes in Computer Science 1552, Springer-Verlag, Berlin. 41-54.

Agrawal, R. and Psaila, G. 1995. 'Active Data Mining'. In Proc. First International Conference on Knowledge Discovery and Data Mining (KDD-95), Montreal, Quebec, Canada. AAAI Press, Menlo Park, California. 3-8.

Bettini, C., Wang, X.S. and Jajodia, S. 1996. 'Testing complex temporal relationships involving multiple granularities and its application to data mining (extended abstract)'. In Proc. Fifteenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, ACM Press. 68-78.

Blum, R.L. 1982. 'Discovery, Confirmation, and Incorporation of Causal Relationships from a Large Time-Oriented Clinical Database: The RX Project'. In Computers and Biomedical Research. 15, 164-187.

Chakrabarti, S., Sarawagi, S. and Dom, B. 1998. 'Mining surprising patterns using temporal description length'. In Proc. Twenty-Fourth International Conference on Very Large databases VLDB'98, New York, NY. Morgan Kaufmann. 606-617.

Chen, X. and Petrounias, I. 1998. 'Language support for temporal data mining'. In Proc. Second European Symposium on Principles of Data Mining and Knowledge Discovery, PKDD'98, Springer-Verlag, Berlin. 282-290.

Chen, X., Petrounias, I. and Heathfield, H. 1998. 'Discovering temporal association rules in temporal databases'. In Proc. International Workshop on Issues and Applications of Database Technology (IADT'98), 312-319.

Imam, I.F. 1994. 'An experimental study of discovery in large temporal databases'. In Proc. Seventh International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems, IEA/AIE '94., 171-180.

Long, J.M., Irani, E.A. and Slagle, J.R. 1991. 'Automating the Discovery of Causal Relationships in a Medical Records Database'. In Knowledge discovery in databases. G. Piatetsky-Shapiro and W.J. Frawley (eds.), AAAI Press/MIT Press, 465-476.

Rainsford, C.P. and Roddick, J.F. 1997. 'An attribute-oriented induction of rules from temporal interval data'. In Proc. Eighth International Database Workshop, Data Mining, Data Warehousing and Client/Server Databases (IDW'97), Hong Kong. Springer Verlag. 108-118.

Saraee, M.H. and Theodoulidis, B. 1995. 'Knowledge discovery in temporal databases'. In Proc. IEE Colloquium on 'Knowledge Discovery in Databases', IEE, London. 1-4.

Wijsen, J. and Meersman, R. 1997. 'On the complexity of mining temporal trends'. In Proc. SIGMOD'97 Workshop on Research Issues on Data Mining and Knowledge Discovery, Tucson, Arizona.

2.3 Discovery of Temporal Patterns

Agrawal, R., Lin, K.-I., Sawhney, H.S. and Shim, K. 1995. 'Fast similarity search in the presence of noise, scaling, and translation in time-series databases'. In *Proc. Twenty-First International Conference on Very Large Data Bases*, San Francisco, CA. Morgan Kaufmann. 490-501.

Agrawal, R. and Srikant, R. 1995. 'Mining sequential patterns'. In *Proc. Eleventh International Conference on Data Engineering*, Taipei, Taiwan. IEEE Computer Society Press. 3-14.

Bayardo Jr, R.J. 1998. 'Efficiently mining long patterns from databases'. In *Proc. ACM SIGMOD Conference on the Management of Data*, Seattle, WA. ACM Press. 85-93.

Berger, G. and Tuzhilin, A. 1998. 'Discovering unexpected patterns in temporal data using temporal logic'. In *Temporal Databases - Research and Practice*. O. Etzion, S. Jajodia and S. Sripada (eds.), Lecture Notes in Computer Science 1399, Springer-Verlag, Berlin. 281-309.

Chen, M.S., Park, J.S. and Yu, P.S. 1996. 'Data mining for path traversal patterns in a web environment'. In *Proc. Sixteenth International Conference on Distributed Computing Systems*, 385-392.

Dietterich, T.G. and Michalski, R.S. 1985. 'Discovering patterns in sequences of events'. *Artif. Intell.* **25**187-232.

Han, J., Gong, W. and Yin, Y. 1998. 'Mining segment-wise periodic patterns in time-related databases'. In *Proc. Fourth International Conference on Knowledge Discovery and Data Mining*, AAAI Press, Menlo Park. 214-218.

Klösgen, W. 1995. 'Deviation and association patterns for subgroup mining in temporal, spatial, and textual data bases'. In *Proc. First International Conference on Rough Sets and Current Trends in Computing, RSCTC'98*, Springer-Verlag, Berlin,. 1-18.

Lin, D.-I. and Kedem, Z.M. 1998. 'Pincer Search: A new algorithm for discovering the maximum frequent set'. In *Proc. International Conference on Extending Database Technology, EDBT'98*, Valencia, Spain. 385-392.

Mannila, H. and Toivonen, H. 1996. 'Discovering generalised episodes using minimal occurences'. In *Proc. Second International Conference on Knowledge Discovery and Data Mining (KDD-96)*, Portland, Oregon. AAAI Press, Menlo Park. 146-151.

Mannila, H., Toivonen, H. and Verkamo, A.I. 1995. 'Discovering frequent episodes in sequences'. In *Proc. First International Conference on Knowledge Discovery and Data Mining (KDD-95)*, Montreal, Quebec, Canada. AAAI Press, Menlo Park, California. 210-215.

Padmanabhan, B. and Tuzhilin, A. 1996. 'Pattern discovery in temporal databases: a temporal logic approach'. In *Proc. Second International Conference on Knowledge Discovery and Data Mining*, Portland, Oregon. AAAI Press.

Pirolli, P., Pitkow, J. and Rao, R. 1996. 'Silk from a Sow's Ear: Extracting usable structures from the web'. In *Proc. CHI'96*, Vancouver.

Spiliopoulou, M. 1998. 'The laborious way from data mining to web mining'. *International Journal on Computer Systems Science and Engineering.* To appear.

Spiliopoulou, M. and Faulstich, L.C. 1998. 'WUM: A tool for web utilization analysis'. In *Proc. EDBT Workshop WebDB'98*, Valencia, Spain. Springer Verlag. Lecture Notes in Computer Science In press.

Srikant, R. and Agrawal, R. 1996. 'Mining sequential patterns: generalisations and performance improvements'. In *Proc. International Conference on Extending Database Technology, EDBT'96*, Avignon, France.

Wade, T.D., Byrns, P.J., Steiner, J.F. and Bondy, J. 1994. 'Finding temporal patterns - a set based approach'. *Artificial Intelligence in Medicine.* (6):263-271.

Wang, K. 1997. 'Discovering patterns from large and dynamic sequential data'. *Intel. Inf. Syst.* **8**8-33.

Wang, K. and Tan, J. 1996. 'Incremental discovery of sequential patterns'. In *Proc. ACM SIGMOD Workshop on Research Issues on Data Mining and Knowledge Discovery*, Montreal, Canada.

Weiss, G.M. and Hirsh, H. 1998. 'Learning to predict rare events in event sequences'. In *Proc. Fourth International Conference on Knowledge Discovery and Data Mining.*, AAAI Press, Menlo Park, CA. 359-363.

Wexelblat, A. 1996. 'An environment for aiding information-browsing tasks'. In *Proc. AAAI Spring Symposium on Acquisition, Learning and Demonstration: Automating Tasks for Users*, Birmingham, England. AAAI Press.

Zaki, M.J. 1997. 'Fast mining of sequential patterns in very large databases'. Technical Report 668. University of Rochester.

Zaki, M.J., Lesh, N. and Ogihara, M. 1998. 'PLANMINE: Sequence mining for plan failures'. In *Proc. Fourth International Conference on KDD*, New York, NY. 369-373.

2.4 Similar Trends in Time Series

The citations in this subsection originate from the domain of time series analysis. They propose efficient algorithms that match a time series to a pattern. This pattern may describe a trend like upward or steep upward move or it may be a time (sub)series itself.

Agrawal, R., Psaila, G., Wimmers, E.L. and Zaot, M. 1995. 'Querying shapes of histories'. In *Proc. Twenty-first International Conference on Very Large Databases (VLDB '95)*, Zurich, Switzerland. Morgan Kaufmann Publishers, Inc. San Francisco, USA. 490-501.

Berndt, D.J. and Clifford, J. 1995. 'Finding patterns in time series: a dynamic programming approach'. In *Advances in Knowledge Discovery and Data Mining*. U.M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, *et al.* (eds.), AAAI Press/ MIT Press, 229-248.

Faloutsos, C., Ranganathan, M. and Manolopoulos, Y. 1994. 'Fast subsequence matching in time-series databases'. In *Proc. ACM SIGMOD Conference on the Management of Data*, Minneapolis. 419-429.

Keogh, E. and Smyth, P. 1997. 'A probabilistic approach to fast pattern matching in time series databases'. In *Proc. Third International Conference on Knowledge Discovery and Data Mining*, Newport Beach, California. AAAI Press, Menlo Park, California. 24-30.

3. SPATIAL DATA MINING

While spatial data mining can be superficially considered as the multi-dimensional equivalent of temporal data mining, in practice the scaling up of dimensions yields only a limited number of useful techniques. Thus, the spatial data mining research to date has generally taken the alternative path of embedded uniquely spatial constructs on top of static techniques; association rules, clustering and characterisation all have their spatial analogues.

One particular problem affecting spatial data mining is the lack of agreement on a set of common underlying data formats. This is hardly surprising given the range of spatial application domains and, while not strictly a data mining research problem, adds a practical hurdle to the implementation of proof-of-concept systems.

The papers below can be categorised in a number of ways; by resulting rule type, by application domain (and by implication the richness of the spatial data relative to the static component) and by volume of data. Given the immaturity of the field at present any categorisation within a short bibliography such as this would be arbitrary and may mask many important contributions.

Bell, D.A., Anand, S.S. and Shapcott, C.M. 1994. 'Data Mining in Spatial Databases'. In *Proc. International Workshop on Spatio-Temporal Databases*, Benicassim, Spain.

Ester, M., Frommelt, A., Kriegel, H.P. and Sander, J. 1998. 'Algorithms for characterization and trend detection in spatial databases'. In *Proc. Fourth International Conference on Knowledge Discovery and Data Mining.*, AAAI Press, Menlo Park. 44-50.

Ester, M., Kriegel, H.-P. and Sander, J. 1997. 'Spatial Data Mining: A Database Approach'. In *Proc. Fifth Symposium on Large Spatial Databases (SSD'97)*, Berlin, Germany. Springer-Verlag. Lecture Notes in Computer Science 48-66.

Ester, M., Kriegel, H.-P., Sander, J. and Xu, X. 1996. 'A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise'. In *Proc. Second International Conference on Knowledge Discovery and Data Mining*, Portland, Oregon. AAAI Press.

Ester, M., Kriegel, H.-P. and Xu, X. 1995. 'Knowledge Discovery in Large Spatial Databases: Focusing Techniques for Efficient Class Identification'. In *Proc. Fourth International Symposium on Large Spatial Databases*, Maine.

Estivill-Castro, V. and Murray, A.T. 1998. 'Discovering associations in spatial data-an efficient medoid based approach'. In *Proc. Second Pacific-Asia Conference on Research and Development in Knowledge Discovery and Data Mining, PAKDD-98*, Springer-Verlag, Berlin. 110-121.

Han, J., Koperski, K. and Stefanovic, N. 1997. 'GeoMiner: A System Prototype for Spatial Data Mining'. In *Proc. ACM SIGMOD Conference on the Management of Data*, Tucson, Arizona.

Kang, I.-S., Kim, T.-W. and Li, K.-J. 1997. 'A Spatial Data Mining Method by Delaunay Triangulation'. In *Proc. Fifth ACM Workshop on Geographic Information Systems*, Las Vegas, Nevada.

Klösgen, W. 1995. 'Deviation and association patterns for subgroup mining in temporal, spatial, and textual data bases'. In *Proc. First International Conference on Rough Sets and Current Trends in Computing, RSCTC'98*, Springer-Verlag, Berlin,. 1-18.

Knorr, E.M. and Ng, R.T. 1996. 'Extraction of Spatial Proximity Patterns by Concept Generalization'. In *Proc. Second International Conference on Knowledge Discovery and Data Mining*, Portland, Oregon. AAAI Press. 347-350.

Knorr, E.M. and Ng, R.T. 1996. 'Finding aggregate proximity relationships and commonalities in spatial data mining'. *IEEE Trans. Knowl. and Data Eng.* 8(6):884-897.

Knorr, E.M., Ng, R.T. and Shilvock, D.L. 1997. 'Finding boundary shape matching relationships in spatial data'. In *Advances in Spatial Databases - Proc. 5th International Symposium, SSD '97*. Springer-Verlag, Berlin, 29-46.

Koperski, K., Adhikary, J. and Han, J. 1996. 'Knowledge Discovery in Spatial Databases: Progress and Challenges'. In *Proc. ACM SIGMOD Workshop on Research Issues on Data Mining and Knowledge Discovery*, Montreal, Canada. 55-70.

Koperski, K. and Han, J. 1995. 'Discovery of Spatial Association Rules in Geographic Information Databases'. In *Proc. Fourth International Symposium on Large Spatial Databases*, Maine. 47-66.

Ng, R.T. 1996. 'Spatial Data Mining: Discovering Knowledge of Clusters from Maps'. In *Proc. ACM SIGMOD Workshop on Research Issues on Data Mining and Knowledge Discovery*, Montreal, Canada.

Ng, R.T. and Han, J. 1994. 'Efficient and effective clustering methods for spatial data mining'. In *Proc. Twentieth International Conference on Very Large Data Bases*, Santiago, Chile.

Popelinsky, L. 1998. 'Knowledge discovery in spatial data by means of ILP'. In *Proc. Second European Symposium on the Principles of Data Mining and Knowledge Discovery, PKDD'98*, Springer-Verlag, Berlin. 185-193.

Shek, E.C., Muntz, R.R., Mesrobian, E. and Ng, K. 1996. 'Scalable Exploratory Data Mining of Distributed Geoscientific Data'. In *Proc. Second International Conference on Knowledge Discovery and Data Mining*, Portland, Oregon. AAAI Press.

Son, E.-J., Kang, I.-S., Kim, T.-W. and Li, K.-J. 1998. 'A spatial data mining method by clustering analysis'. In *Proc. Sixth International Symposium on Advances in Geographic Information Systems, GIS'98*, 157-158.

Zhang, T., Ramakrishnan, R. and Livny, M. 1996. 'BIRCH: An Efficient Clustering Method for Very Large Databases'. In *Proc. ACM SIGMOD Workshop on Research Issues on Data Mining and Knowledge Discovery*, Montreal, Canada.

Zhou, X., Truffet, D. and Han, J. 1999. 'Efficient Polygon Amalgamation Methods for Spatial OLAP and Spatial Data Mining'. In Proc. 6th International Symposium on Spatial Databases (SSD'99), Hong Kong. To appear.

4. SPATIO-TEMPORAL DATA MINING

Research accommodating both spatial and temporal data mining is sparse at present; however, two directions can be identified. Firstly, the embedding of a temporal awareness in spatial systems, and secondly, the accommodation of space into temporal data mining systems. To date, the former approach has been the more popular, in part because of the relative maturity of geographic information systems and the availability of time-stamped snapshots of geographic/spatial test data.

This research area is expected to grow considerably in the next few years with a number of events looking at this issue. As well as the existing KDD conferences, four such events include:

- Dagstuhl Seminar on Integrating Spatial and Temporal Databases, Wadern, Germany, November 1998, Dagstuhl Report No.\ 228.
- NCGIA Varenius Workshop on Discovering Geographic Knowledge in Data-Rich Environments, Redmond WA, USA, March 1999.
- DEXA Workshop on Spatio-Temporal Data Models and Languages, Florence, Italy, August 1999.
- VLDB Workshop On Spatio-Temporal Database Management, Edinburgh, Scotland, September 1999.

Abraham, T. and Roddick, J.F. 1997. 'Discovering meta-rules in mining temporal and spatio-temporal data'. In *Proc. Eighth International Database Workshop, Data Mining, Data Warehousing and Client/Server Databases (IDW'97)*, Hong Kong. Springer-Verlag. 30-41.

Abraham, T. and Roddick, J.F. 1997. 'Research issues in spatio-temporal knowledge discovery'. In *Proc. SIGMOD'97 Workshop on Data Mining*, Arizona, USA. ACM Press. 85.

Abraham, T. and Roddick, J.F. 1998. 'Opportunities for knowledge discovery in spatio-temporal information systems'. *Aust. J. Inf. Syst.* 5(2):3-12.

Mesrobian, E., Muntz, R., Shek, E., Nittel, S., La-Rouche, M., Kriguer, M., Mechoso, C., Farrara, J., Stolorz, P. and Nakamura, H. 1996. 'Mining geophysical data for knowledge'. *IEEE Expert.* **11**(5):34-44.

Mesrobian, E., Muntz, R., Shek, E., Santos, J.R., Yi, J., Ng, K., Chien, S.Y., Mechoso, C., Farrara, J., Stolorz, P. and Nakamura, H. 1995. 'Exploratory data mining and analysis using CONQUEST'. In *Proc. IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, IEEE, New York. 281-286.

Stolorz, P. and Dean, C. 1996. 'Quakefinder: A Scalable Data Mining System for Detecting Earthquakes from Space'. In *Proc. Second International Conference on Knowledge Discovery and Data Mining (KDD96)*, Portland, Oregon. AAAI Press, Menlo Park, California. 208-213.

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. and Farrara, J.D. 1995. 'Fast Spatio-Temporal Data Mining of Large Geophysical Sets'. In Proc. First International Conference on Knowledge Discovery and Data Mining, Montreal, Canada. AAAI Press. 300-305.

About the authors:

John Roddick is currently an Associate Professor in the School of Computer and Information Science at the University of South Australia. Prior to joining the University nine years ago, he worked as a database administrator and consultant at the University of Tasmania, Australia, and at Mobil Data Services and Sun Alliance Insurance in the UK. His technical interests include data mining, schema versioning and data warehouse construction. He holds a PhD from La Trobe University, an MSc from Deakin University and a BSc(Eng)(Hons) from Imperial College, London. He is a member of the Australian Computer Society, the IEEE Computer Society and the Association for Computing Machinery.

Myra Spiliopoulou received her BSc degree in mathematics and the PhD degree in computer science from the University of Athens, Greece, in 1986 and 1992, respectively. Between 1987-1994, she worked as a research assistant in the Department of Informatics, University of Athens, and was involved in national and European projects on parallel database query optimization, hypermedia and multimedia modelling and querying, and on computer-aided education.

Dr Spiliopoulou is assistant professor with the Institute of Information Systems in the Economics Faculty of the Humboldt-University in Berlin since 1994. Her research interests include sequence analysis, temporal mining, web usage analysis and user profiling. She works both on the algorithmical aspects of knowledge discovery and on its practical applications in areas such as electronic commerce and education. She is a member of the IEEE Computer Society and of the Association for Computing Machinery.