An Overview of Temporal Data Mining

Weiqiang Lin
Department of Computing
I.C.S., Macquarie University
Sydney, NSW 2109, Australia
wlin@ics.mq.edu.au

Mehmet A. Orgun
Department of Computing
I.C.S., Macquarie University
Sydney, NSW 2109, Australia
mehmet@ics.mq.edu.au

Graham J. Williams
CSIRO Data Mining
GPO Box 664
Canberra ACT 2601, Australia
Graham.Williams@csiro.au

ABSTRACT

Temporal Data Mining is a rapidly evolving area of research that is at the intersection of several disciplines, including statistics, temporal pattern recognition, temporal databases, optimisation, visualisation, high-performance computing, and parallel computing. This paper is first intended to serve as an overview of the temporal data mining in research and applications.

1. INTRODUCTION

Temporal Data Mining is a rapidly evolving area of research that is at the intersection of several disciplines, including statistics (e.g., time series analysis), temporal pattern recognition, temporal databases, optimisation, visualisation, high-performance computing, and parallel computing. This paper is intended to serve as an overview of the temporal data mining in research and applications. In addition to providing a general overview, we motivate the importance of temporal data mining problems within Knowledge Discovery in Temporal Databases (KDTD) which include formulations of the basic categories of temporal data mining methods, models, techniques and some other related areas.
The paper is structured as follows. Section 2 discusses the definitions and tasks of temporal data mining. Section 3 discusses the issues on temporal data mining techniques. Section 4 discusses two major problems of temporal data mining, those of similarity and periodicity. Section 5 provides an overview of time series temporal data mining. Section 6 moves onto a discussion of several important challenges in temporal data mining and outlines our general distribution theory for answering some those challenges. The last section concludes the paper with a brief summary.

2. DEFINITION AND TASKS OF TEMPORAL DATA MINING

The temporal data mining component of the KDTD process is concerned with the algorithmic means by which temporal patterns are extracted and enumerated from temporal data. Some problems for temporal data mining in temporal databases include questions such as: How can we provide access to temporal data when the user does not know how to describe the goal in terms of a specific query? How can we find all the time related information and understand a large temporal data set? and so on.

2.1 Definition and Aims

In this section, we first give basic definitions and aims of Temporal Data Mining. The definition of Temporal Data Mining is as follows:

Definition 1. *Temporal Data Mining is a single step in the process of Knowledge Discovery in Temporal Databases that enumerates structures (temporal patterns or models) over the temporal data, and any algorithm that enumerates temporal patterns from, or fits models to, temporal data is a Temporal Data Mining Algorithm.*

Basically temporal data mining is concerned with the analysis of temporal data and for finding temporal patterns and regularities in sets of temporal data. Also temporal data mining techniques allow for the possibility of computer-driven, automatic exploration of the data. Temporal data mining has led to a new way of interacting with a temporal database: specifying queries at a much more abstract level than say, Temporal Structured Query Language (TSQL) permits (e.g., [17], [16]). It also facilitates data exploration for problems that, due to multiple and multi-dimensionality, would otherwise be very difficult to explore by humans, regardless of use of, or efficiency issues with, TSQL.
Temporal data mining tends to work from the data up and the best known techniques are those developed with an orientation towards large volumes of time related data, making use of as much of the collected temporal data as possible to arrive at reliable conclusions. The analysis process starts with a set of temporal data, uses a methodology to develop an optimal representation of the structure of the data during which time knowledge is acquired. Once Temporal knowledge has been acquired, this process can be extended to a larger set of the data working on the assumption that the larger data set has a structure similar to the sample data.

2.2 Temporal Data Mining Tasks

A relevant and important question is how to apply data mining techniques on a temporal database. According to techniques of data mining and theory of statistical time series analysis, the theory of temporal data mining may involve the following areas of investigation since a general theory for this purpose is yet to be developed:

1. Temporal data mining tasks include:
   - Temporal data characterization and comparison,
   - Temporal clustering analysis,

The Australasian Data Mining Workshop

- Temporal classification,
- Temporal association rules,
- Temporal pattern analysis, and
- Temporal prediction and trend analysis.

2. A new temporal data model (supporting time granularity and time-hierarchies) may need to be developed based on:

   - Temporal data structures, and
   - Temporal semantics.

3. A new temporal data mining concept may need to be developed based on the following issues:

   - the task of temporal data mining can be seen as a problem of extracting an interesting part of the logical theory of a model, and
   - the theory of a model may be formulated in a logical formalism able to express quantitative knowledge and approximate truth.

In addition, temporal data mining needs to include an investigation of tightly related issues such as temporal data warehousing, temporal OLAP, computing temporal measurements, and so on.

3. TEMPORAL DATA MINING TECHNIQUES

A common form of a temporal data mining technique is rule (or functions) discovery. Various types of temporal functions can be learnt, depending upon the application domain. Also, temporal functions (or rules) can be constructed in various ways. They are commonly derived by one of the two basic approaches, bottom-up or top-down induction.

3.1 Classification in Temporal Data Mining

The basic goal of temporal classification is to predict temporally related fields in a temporal database based on other fields. The problem in general is cast as determining the most likely value of the temporal variable being predicted given the other fields, the training data in which the target variable is given for each observation, and a set of assumptions representing one's prior knowledge of the problem. Temporal classification techniques are also related to the difficult problem of density estimation.

In recent years, a lot of the work has been done in non-temporal classification areas by using “Statistical Approaches to Predictive Modelling”. Some techniques have been established for estimating a categorical variable, e.g., [26; 5; 20]: kernel density estimators [20; 11] and K-nearest-neighbor method [20]. These techniques are based upon the theory of statistics. Some other techniques such as in [7; 8; 6] are based upon the theory of databases. Temporal classification techniques have not been paid much attention so far. In recent years, the main idea in temporal classification is the straightforward use of sampling techniques within time series methods (distribution) to build up a model for temporal sequences.

3.2 Temporal Cluster Analysis

Temporal clustering according to similarity is a concept which appears in many disciplines, so there are two basic approaches to analyze it. One is the measure of temporal similarity approach and the other is called temporal optimal partition approach.

In temporal data analysis, many temporal data mining applications make use of clustering according to similarity and optimization of temporal set functions. If the number of clusters is given, then clustering techniques can be divided into three classes: (1) Metric-distance based technique, (2) Model-based technique and (3) Partition-based technique. These techniques can be used occasionally in combination, such as Probability-based vs. Distance-based clustering analysis. If the number of clusters is not given, then we can use Non-Hierarchical Clustering Algorithms to find their k.

In recent years, temporal clustering techniques have been developed for temporal data mining, e.g., [23]. Some studies have been done by using EM algorithm and Monte-Carlo cross validation approach (e.g.,[12; 22; 13]).

3.3 Induction

A temporal database is a store of temporally related information but more important is the information which can be inferred from it([3; 4]. There are two main inference techniques: temporal deduction and temporal induction.

1. Temporal deduction is a technique (e.g., in [24] to infer the information that is a temporal logical consequence of the information in the temporal database.
2. Temporal induction can be described as a technique (e.g., in [25]) to infer temporal information that is generalised from the temporal database. Induction has been used in the following ways within data mining: 1) Decision Trees and 2) Rule Induction.

4. TWO FUNDAMENTAL TEMPORAL DATA MINING PROBLEMS

In recent years, two kinds of fundamental problems have been studied in temporal data mining area. One is the Similarity Problem which is to find a time sequence (or TDB) similar to a given sequence (or query) or to find all pairs of similar sequences. The other is the Periodical Problem which is to find periodic patterns in TDB.

4.1 Similarity Problems

In temporal data mining applications, it is often necessary to search within a temporal sequence database (e.g: TDB) for those sequences that are similar to a given query sequence. Such problems are often called Similarity Search Problem. This kind of a problem involves search on multiple and multidimensional time series sets in TDBs to find out how many series are similar to one another. It is one of the most important and growing problems in Temporal Data Mining. In recent years, we still lack a standard definition and standard theory for similarity problems in TDB.

Temporal data mining techniques can be applied in similarity problems. The main steps for solving the similarity problem are as follows:

- define similarity: allows us to find similarities between sequences with different scaling factors and baseline values.

- choose a query sequence: allows us to find what we want to know from large sequences (TDB) (e.g, character, classification)

- processing algorithm for TDB: allows us to apply some statistical methods (e.g, transformation, wavelet analysis) to TDB (e.g, remove the noisy data, interpolate the missing data).

- processing an approximate algorithm: allows us to build up a classcification scheme for the TBD according to the definition of similarity by using some data mining techniques (e.g, visualisation).

The result of the Similarity Problem search in TDB can be used for temporal association, prediction, etc.

4.2 Periodical Problems

The periodicity problem is the problem of finding periodic patterns or, cyclicity occurring in time-related databases (TDB). The problem is related to two concepts: *pattern* and *interval*. In any selected sequence of TDB, we are interested in finding patterns which repeat over time and their recurring intervals (period), or finding the repeating patterns of a sequence (or TDB) as well as the interval which corresponds to the pattern period. For solving a Periodical Problem in TDB, the main steps are as follows:

- determining some definitions of the concept of a period under some assumptions: this step allows us to know what kind of a periodicity search we want to perform from TDB.

- building up a set of algorithms: this step allows us to use properties of periodic time series for finding periodic patterns from a subset of TDB by using algorithms.

- processing simulation algorithms: this step allows us find patterns from whole TDB by the algorithms.

A lot of techniques have been involved in these kind of problems by using pure mathematical analysis such as function analysis, data distribution analysis and so on, e.g.,[9].

4.3 Discussion

In a time-series TDB, sometimes similarity and periodical search problems are difficult even when there are many existing methods, but most of the methods are either inapplicable or prohibitively expensive. There is also another difficult problem: how we can combine multiple-level similarity or periodical search in a multiple-level model? With the reference cube structure, such difficult problems can be solved by extending the methods mentioned in previous subsections, but the problem of combining multiple-level similarity and periodicity in a multiple-level model is still unsolved. Also, more sophisticated techniques need to be developed to reduce memory work-space.
In fact, similarity and periodical search problems can be combined into the problem of finding interesting sequential patterns in TDBs. In recent years, some new algorithms have been developed for "fast mining of sequential patterns in large TDBs":

- generalized sequential pattern (GSP) algorithm: it essentially performs a level-wise or breadth-first search of the sequence lattice spanned by the subsequence relation,

- sequential pattern discovery using equivalence classes (SPADE) algorithm: it decomposes the original problem into smaller sub-problems using equivalence classes on frequent sequences [15].

With any new algorithm, there is one important question that has often been asked: How can we implement the new algorithm directly on top of a Time-series TDB?

5. TIME SERIES TEMPORAL DATA MINING

Statistics has been an important tool for data analysis for a long time. For example, Bayesian inference is the most extensively studied statistical method for knowledge discovery (e.g, [2], [10], [18]) and Markov Model, Hidden Markov Model (e.g., [14; **?**]) also have made their way into temporal knowledge discovery process.
Time series is a record of the values of any fluctuating quantity measured at different points of time. One characteristic feature which distinguishes time series data from other types of data is that, in general, the values of the series at different time instants will be correlated[1]. Application of time series analysis techniques in temporal data mining is often called Time Series Data Mining. A great deal of work has been done into identifying, gathering, cleaning, and labeling the data, into specifying the questions to be asked of it, and into finding the right way to view it to discover useful temporal patterns.
Time series analysis method has been applied into following major categories in temporal data mining:

1. Representation of Temporal Sequence: This refers to the representation of data before actual temporal data mining techniques take place. There are two major methods:

   - General representation of data: representation of data into time series data in either continuous or discontinuous, linear/non-linear models, stationary/non-stationary models and distribution models (e.g., Time domain representation and Time series model representation).

   - General transformation of representation of data: representation of data into time series data in either continuous or discontinuous transformation (e.g., Fourier transformation, Wavelet transformation and Discretization transformation).

2. Measure of Temporal Sequence: measuring temporal charactersistic element in given definitions of similarity and/or periodicity in a temporal sequence (or, two subsequence in a temporal sequence) or between temporal sequences. There are two methods:

[1]Time Analysis Theory can be found in any standard textbook of time series analysis, e.g., [1].

- Characteristic distance measuring in time domain: measuring distance between temporal charactersistics in either continuous or discontinuous time domain (e.g., Euclidean squared distance function).
- Characteristic distance measuring in other than time domain: measuring distance between temporal charactersistics in either continuous or discontinuous domain other than time (e.g., distance function between two distributions).

3. Prediction of Temporal Sequence: the main goal of prediction is to predict some fields in a database based on Time domain. The techniques can be classified into two models.
   - Temporal classification models: the basic goal is to predict the most likely state of a categorical variable (the class) in Time domain.
   - Temporal regression models: the basic goal is to predict a numeric variable in a set by using different transformations (e.g, linear or non-linear) on databases to find temporal information (or, patterns) of the different (or the same) categorical data sets (class).

Recently, there are various results to date on discovering temporal information which have offered forums to explore the temporal data mining progress and future work concerning temporal data mining. But the general theory and general method of temporal data analysis of discovering temporal patterns for temporal sequence data analysis are not well known.

6. CHALLENGES AND RESEARCH DIRECTIONS

Recent advances in data collection and storage technologies have made it possible for companies, administrative agencies and scientific laboratories to keep vast amounts of temporal data relating to their activities. Data mining refers to such an activity to make automatic extraction of different levels of knowledge from data feasible. One of the main unresolved problems, often called General Analysis Method of Temporal Data Mining, that arise during the data mining process is treating data that contains temporal information.

6.1 Challenge Questions

Data mining is a step in the knowledge discovery in databases, although successful data mining applications continue to appear but the fundamental problems are still as difficult as they have been for the past decade. One such difficult and fundamental problem is the development of a general data mining analysis theory. Temporal data mining researchers have paid some attention to this problem but results still remain in their infancy. One of the important roots in data mining analysis is statistical analysis theory. The general temporal data mining analysis theory includes two important analysis methods:

- Data structural temporal knowledge analysis method: This method involves the discovery of data prior temporal knowledge, and the exploitation of the knowledge into a data analysis model to establish the link between the present temporal knowledge and the future temporal knowledge.
- Data temporal measure analysis method: This method involves the transformation of initial data temporal domain (or space) into another domain (or space), then the use of this new domain to represent the original temporal data.

The techniques involved in the above two methods can be divided into following classes:

1. **Temporal data clustering:** temporal clustering targets separating the temporal data into subsets that are similar to each other. There are two fundamental problems of temporal clustering:
   - To define a meaningful similarity measure, and,
   - To choose the number of temporal clusters(if we do not know the cluster numbers).
2. **Temporal data prediction:** the goal of temporal prediction is to predict some fields based on other temporal fields. Temporal data prediction also involves using prior temporal patterns (or, models, knowledge) for finding the data attributes relevant to the attribute of interest.
3. **Temporal data summarization:** the purpose of temporal data summarization is to describe a subset of temporal data by representing extracted temporal information in a model or, in rules or in patterns. It provides a compact description for a temporal dataset. It could also involve a logic language such as temporal logic, fuzzy logic and so on.
4. **Temporal data dependency:** Temporal dependency modelling describes time dependencies among data and/or temporal attributes of data. There are two dependency models: qualitative and quantitative. The qualitative dependency models specify temporal variables (e.g., time gap) that are locally dependent on a given state-space $\mathcal{S}$. The quantitative dependency models specify the value dependencies (e.g., using numerical scale) in a statistical space $\mathcal{P}$.

6.2 Some Answers of the Challenge Questions

During the past few years, we have proposed a formal framework for the definitions and general hidden distribution theory of temporal data mining. We have also investigated applications in temporal clustering, temporal classification and temporal feature selection for temporal data mining. The major work we have done in answering the temporal data mining challenge questions are:

- We have established a General Hidden Distribution-based Analysis Theory for temporal data mining. The general mining analysis theory is based on the statistical analysis method but traditional statistical assumptions only come from the data itself. There are two important concepts in the theory: 1) data qualitative set, data quantitative set and 2) data hidden conditional distribution function. The data qualitative set is the set to decide the data moving structure such as data periodicity and similarity. In other words, data qualitative set is a base of the data. The data quantitative set is the set to decide the numerical range of the data moving structure. The data hidden conditional distribution function is built on the characteristics of data qualitative and data quantitative sets. Another feature of the general mining analysis method is that we can use (extension of) all existing statistical analysis methods and techniques for mining temporal patterns.

- We have proposed an algorithm which is called The Additive Distributional Recursion Algorithm (ADRA) in General Hidden Distribution-based Analysis Theory for building up temporal data models. The algorithm uses the sieve method[2] to discover temporal distribution function (models, pattern).

- We have extended a normal measure method to a new Temporal Measure Method, which is called Time-gap Measure Method. The new measure method has brought "time length" (which is between temporal events) or "time interval" (which is within a temporal event) into a time point (or, time value) variable. After a temporal sequence is transformed, it can be measured in both state-space $\mathcal{S}$ and probability space $\mathcal{P}$. The time-gap is used as a temporal variable in time distribution function $f(t_v)$ or temporal variable functional equations embedded in temporal models of the sequence.

- We have extended and built up a new application of fundamental mathematics techniques for dealing with large temporal datasets, massive temporal datasets and distributed temporal datasets. The new application is called Temporal Sequence Set-Chains (A special case of the Temporal Set-Chains is a Markov Set-Chains). The key issue in temporal sequence set-chains is the use of stochastic matrices of samples to build up a moving kernel distribution. The temporal sequence set-chains sequence can be used for mining a large temporal sequence, massive temporal sequence and distributed temporal sequence such as Web temporal data sequence (e.g., Web content sequence, Web usage sequence and Web strutural sequence).

- We have proposed a framework of Temporal Clustering method for discovering temporal patterns. In our temporal clustering method, there are three stages of temporal data mining in temporal clustering analysis: 1) the input stage: what an appropriate measure of similarity to use, 2) the algorithm stage: what types of algorithms to use, and 3) the output stage: assessing and interpreting the results of cluster analysis. In the second stage, we also proposed a framework of Distribution-based Temporal Clustering Algorithm. The algorithm is based on our general analysis method.

- We have proposed a framework of Temporal Classification. This temporal classification is generated by our Temporal Clustering method. According to our general analysis theory for Temporal Sequence Mining and its application in temporal clustering, there are also the following three steps for constructing a temporal classification: 1) Provide a definition of temporal classification, 2) Define a distribution distance function and 3) Provide the weighting of temporal objects for changing their class membership. For large numbers of classification, we proposed a *discriminant coordinates of time gap distribution* to deal with such kinds of problems.

- We have proposed a framework of Temporal Feature Selection for discovering Temporal patterns. There are three steps for feature selection in the temporal sequence. The first step of the framework employs a distance measure function on time-gap distributions between temporal events for discovering structural temporal features. In this step, only rough shapes of patterns are decided. The temporal features are grouped into temporal classifications by employing a distribution distance measure. In the second step, the degree of similarity and periodicity between the extracted features are measured based on the data value distribution models. The third step of the framework consists of a hybrid model for selecting global features based on the results of the first two steps.

- We have established the main steps of applying our general temporal data mining theory to real world datasets with different methods and models. There are three steps of applications of our general analysis for discovering knowledge from a temporal sequence: 1) preprocessing data analysis including solving data problems and transforming data from its original form into its quantitative set and qualitative set, 2) temporal pattern searching including qualitative-based pattern searching, quantitative-based pattern searching and discovering global temporal patterns (models), and 3) the interpretation of the global temporal patterns (models) and future prediction.

[2]The sieve method is an important method in number theory.

6.3 Future Research Directions

As we mentioned earlier temporal data mining and knowledge discovery have emerged as fundamental research areas with important applications in science, medicine and business. In this section, we describe some of the major directions of research from recent general analysis theory of temporal data mining research:

1. An extension of this temporal sequence measure method to general temporal points (e.g., temporal interval-based gap function) allowing an arbitrary interval between temporal points may lead to a very powerful temporal sequence transformation method.

2. An extension of the notion of Temporal Sequence Set-Chains on different temporal variables, or different components of a temporal variable, can be applied to deal with following problems of temporal data mining:

   - the number of temporally related attributes of each observation increases,
   - the number of temporally related observations increases, and
   - the number of temporally related distribution functions increases.

3. An important extension of the general temporal mining theory is the development of distributed temporal data mining algorithms.

4. In applications of temporal data mining, all new temporal data mining theories, methods and techniques should be developed on/with privacy and security models and protocols appropriate for temporal data mining.

5. In general data mining theory, we may need to develop fundamental mathematical techniques of fuzzy methods for mining purposes (e.g., temporal fuzzy clustering and algorithms, temporal fuzzy association rules and new types of temporal databases.).

7. CONCLUDING REMARKS

Temporal data mining is a very fast expanding field with many new research results reported and many new temporal data mining analysis methods or prototypes developed recently. Some articles of overview of temporal data mining have discussed in different frameworks for coveing research and application in temporal data mining. In [19], for example, Roddick and Spiliopoulou have presented a comprehensive overview of techniques for the mining of temporal data.

In this report we have provided an overview of the temporal data mining process and some background to Temporal Data Mining. Also we discussed a difficult and fundamental problem, a general analysis theory of temporal data mining and provided some answers to the problem. This leads into a discussion on why there was a need for Temporal Data Mining in industry, which has been a major factor in the efforts that have gone into building the present generation of Temporal Data Mining Systems. We have presented a number of areas which are related to Temporal Data Mining in their objectives and compared and contrasted these technologies with Temporal Data Mining.

Acknowledgements

This research has been supported in part by an Australian Research Council (ARC) grant and a Macquarie University Research Grant (MURG).

8. REFERENCES

[1] D. Brillinger, editor. *Time Series: Data Analysis and Theory*. Holt, Rinehart and Winston, New York, 1975.

[2] P. Cheeseman and J. Stutz. Bayesian classification (AUTOCLASS): Theory and results. In U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, editors, *Advances in Knowledge Discovery and Data Mining*. AAAI Press / MIT Press, 1995.

[3] T. Fulton, S. Salzberg, S. Kasif, and D. Waltz. Local induction of decision trees: Towards interactive data mining. In Simoudis et al. [21], page 14.

[4] B. R. Gaines and P. Compton. Induction of meta-knowledge about knowledge discovery. *IEEE Trans. On Knowledge And Data Engineering*, 5:990–992, 1993.

[5] C. Glymour, D. Madigan, D. Pregibon, and P. Smyth. Statistical inference and data mining. *Communications of the ACM*, 39(11):35–41, Nov. 1996.

[6] F. H. Grupe and M. M. Owrang. Data-base mining - discovering new knowledge and competitive advantage. *Information Systems Management*, 12:26–31, 1995.

[7] J. Han, Y. Cai, and N. Cercone. Knowledge discovery in databases: An attribute-oriented approach. In *Proceedings of the 18th VLDB Conference*, pages 547–559, Vancouver, British Columbia, Canada, Aug. 1992.

[8] J. W. Han, Y. D. Cai, and N. Cercone. Data-driven discovery of quantitative rules in relational databases. *Ieee Trans. On Knowledge And Data Engineering*, 5:29–40, Feburary 1993.

[9] J. W. Han, Y. Yin, and G.Dong. Efficient mining of partial periodic patterns in time series database. *IEEE Trans. On Knowledge And Data Engineering*, 1998.

[10] D. Heckerman, H. Mannila, D. Pregibon, and R. Uthurusamy, editors. *Learning bayesian networks: the combineation of knowledge and statistical data*. AAAI Press, 1994.

[11] D. Heckerman, H. Mannila, D. Pregibon, and R. Uthurusamy, editors. *Proceedings of the Third International Conference on Knowledge Discovery and Data Mining (KDD-97)*. AAAI Press, 1997.

[12] E. Keogh and P. Smyth. A probabilistic approach to fast pattern matching in time series databases. page 126.

[13] A. Ketterlin. Clustering sequences of complex objects. In Heckerman et al. [11], page 215.

[14] C. Li and G. Biswas. Temporal pattern generation using hidden markov model based unsuperised classifcation. In *Proc. of IDA-99*, pages 245–256, 1999.

[15] M.J.Zaki. Fast mining of sequential patterns in very large databases. *Uni. of Rochester Technical report*, 1997.

[16] S. a. O.Etzion, editor. *Temporal databases: Research and Practice*. Springer-Verlag,LNCS1399, 1998.

[17] B. Padmanabhan and A. Tuzhilin. Pattern discovery in temporal databases: A temporal logic approach. In Simoudis et al. [21], page 351.

[18] P.sprites, C.Glymour, and R.Scheines. *Causation, Prediction and Search*. Springer-Verlag, 1993.

[19] J. Roddick and M. Spiliopoulou. A survey of temporal knowledge discovery paradigms and methods. *IEEE Transactions on Knowledge and Data Engineering, 2002*.

[20] R.O.Duda and P. Hart. *Pattern classification and scene analysis*. John Wiley and Sons, 1973.

[21] E. Simoudis, J. W. Han, and U. Fayyad, editors. *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96)*. AAAI Press, 1996.

[22] P. Smyth. Clustering using monte carlo cross-validation. In Simoudis et al. [21], page 126.

[23] T.Oates. Identifying distinctive subsequences in multivariate time series by clustering. In *5th International Conference on Knowledge Discovery Data Mining*, pages 322–326, 1999.

[24] J. D. Ullman and C. Zaniolo. Deductive databases: achievements and future directions. *SIGMOD Record (ACM Special Interest Group on Management of Data)*, 19(4):75–82, Dec. 1990.

[25] D. Urpani, X. Wu, and J. Sykes. RITIO - rule induction two in one. In Simoudis et al. [21], page 339.

[26] P. Usama Fayyad and O.L.Mangasarian. Data mining: Overview and optimization opportunities. *INFORMS*, Special issue on Data Mining, 1998.