

Preface

1. Temporal Data Mining

Many real-world applications deal with huge amounts of temporal data. Examples include alarms/events and performance measurements generated by distributed computer systems and by telecommunication networks, the web server logs, online transaction logs, financial data, and sensor data collected from sensor networks. Conventionally, temporal data is classified to either categorical event streams or numerical time series and both types have been intensively studied in data mining and statistics. However, several previously less emphasized aspects of temporal data have proven their importance in emerging applications and posed several challenges calling for more research.

The setting of traditional temporal data analysis is to apply one algorithm on a static, regular and relatively small temporal data set. Many practitioners found existing analysis methods inadequate for their real-world data. Many struggle to transform the data in order to apply existing methods or even to reduce the original problems to better studied ones; either ways induce in more preprocessing effort, more artificial parameters and less interpretable results. We believe these new aspects of temporal data deserve theories and algorithms of their own. Some of these new aspects are:

- Irregularity—Many types of numerical temporal data are not equally paced.
- Asynchronusness— In distributed computing environments like sensor networks, data from different sources tend to be not aligned and hence can not apply synchronous methods.
- Distributed analysis--- A trend in temporal data analysis is to perform data filtering, transformation and analysis as close as possible to the data sources to avoid the prohibitive amount of data being transmitted and analyzed. This new computing paradigm calls for a new theoretical foundation.
- Streaming data--- It is often desirable to analyze temporal data in (near-) real time for monitoring system health, detecting anomaly, alerting exceptional conditions, diagnosing problem, and performing actions including feedback control and self-healing of systems.
- Heterogeneous data types --- It is very common that temporal data is partly categorical events and partly numerical time series. It remains to be an interesting challenging to best analyze all possible data in a uniform way.
- Huge volume— the stream of data can be huge for a long, continuous observation period. Many types of measurements can be obtained from a large number of data sources. This requires designing scalable solutions in analyzing a large volume of temporal data, in terms of both the large number of data points and the large number of types of measurements.

Driven by the new aspects of temporal data, several fundamental problems need to be revisited. Just to name a few of them:

- Prediction
- Correlation

- Regression
- Benchmarking
- Periodic Pattern Mining
- Temporal Association Finding
- Sequential Event Patterns
- Causality Analysis
- Threshold selection
- Frequency Analysis
- Anomaly Detection
- Clustering
- Classification

2. The Workshop

The main objective of the “**Temporal Data Mining: Algorithms, Theory and Applications**” workshop is to bring together researchers from both industry and academia with different backgrounds: data mining, machine learning, database, statistical analysis, and application knowledge to propose new ideas, identify promising technologies and address the aforementioned technical challenges. The homepage for the workshop is <http://www.cs.rochester.edu/~taoli/workshop/>.

The workshop program included 8 technical papers selected by the program committee. It also included invited talks by Dr. Sheng Ma from IBM research and Dr. Oscar Kipersztok from Boeing Research.

We are very indebted to all program committee members who helped us organize the workshop and reviewed the papers very carefully. We would also like to thank all the authors who submitted their papers to the workshop; they provided us with an excellent workshop program.

November 2004

Workshop Co-chairs

Sheng Ma
Tao Li
Chang-shing Perng