The Third SIGKDD Workshop on Mining Temporal and Sequential Data (KDD/TDM 2004)

K. P. Unnikrishnan*
General Motors R & D Center
k.unnikrishnan@gm.com

Ramasamy Uthurusamy
General Motors IS & S
samy@gm.com

Jiawei Han†
University of Illinois
hanj@cs.uiuc.edu

ABSTRACT

In this short report, we provide a summary of the results, issues, and research directions on mining temporal and sequential data, discussed in TDM-2004, held in conjunction with the 10-th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004) on August 22, 2004 in Seattle, Washington, U.S.A.

1. WORKSHOP SUMMARY

This workshop is the third on the topic of Temporal Data Mining following the two successful workshops held at KDD-2001 and KDD-2002. This year's title "Workshop on Mining Temporal and Sequential Data" reflects our initiative to broaden the scope.

Though much of the data in large databases either have explicit or implicit timing/ordering information, many of the successful KDD applications still involve finding static relationships from these databases. Examples of temporal/sequential data include financial data, data from manufacturing plants, data obtained from scientific and engineering processes, sensor data, stream data, weather and geophysical data, traffic-flow data from physical and electronic networks, multi-media and speech data, genomic and other bio-informatics data, brain imaging data, etc. Many of the systems we treat as static, whether the world-wide-web or the human brain, are truly dynamic. Mining of data from these systems would be much easier if the temporal/dynamic nature of these systems is directly addressed.

The goal of this workshop is to bring together researchers and practitioners from multi-disciplinary fields, including temporal data mining, time-series analysis, statistics, sequential pattern analysis, Web analysis, biomedical research, and industrial applications, critically evaluate the need for mining temporal and sequential data, and identify promising technologies and methodologies for making real-time mining of these data possible.

Based on the above motivation, submissions of the papers on the following topics are invited.

- Real-time mining of temporal and sequential data streams
- Data models, techniques and methodologies for mining temporal and sequential data
- Scalability of algorithms for mining temporal and sequential data
- Parallel and distributed mining of temporal and sequential data
- Mining of temporal and sequential data on the web
- Visualization techniques for mining temporal and sequential data

The workshop received a good number of submissions. Based on the critical evaluation by the program committee, we have finally compiled an interesting final workshop program, which consists of 8 selected research papers for workshop presentation, 7 selected papers for poster presentation, one invited talk on "*Data Mining for Financial Time Series: Learning to Trade via Direct Reinforcement*" by John Moody (International Computer Science Institute, Berkeley), and one final panel discussion on "*Where should TDM go?*" coordinated by Unnikrishnan, with the following Panelists: Heikki Mannila, Doug Martin, Vipin Kumar, Eamonn Keogh, Jiawei Han, and Ramasamy Uthurusamy.

The rich program of the workshop attracted good attendance, with lively presentations, posters, and heated discussions. Based on the workshop discussion, we have the following conclusion related to this research theme: Good progress has been made in recent years on time-series similarity search, mining sequential, temporal, spatiotemporal, time-series, and stream data, and their applications. However, this is still an important, dynamic and fast growing subfield in data mining since new types of time- and sequence-related data sets related to sensor and stream data, biomedical data, and spatiotemporal data, as well as new application demands have been the driving force for the fast progress in this field. We predict the dynamic growth and flourishing research on this theme in the years to come.

2. ACKNOWLEDGEMENTS

We would like to thank the workshop program committee members for their service, listed as follows: Rakesh Agrawal, Vasant Dhar, Christos Faloutsos, Usama Fayyad, Johannes Gehrke, Joydeep Ghosh, C. Lee Giles, Eamonn Keogh, Vipin Kumar, Heikki Mannila, R. Douglas Martin, Raymond T. Ng, Arlindo Oliveira, Jian Pei, Gregory Piatetsky-Shapiro, John Roddick, P. S. Sastry, Padhraic Smyth, Ramakrishnan Srikant, and Benjamin Wah.

* We would like to thank General Motors Corporation for help with printing of the workshop notes.

† Work supported in part by the U.S. National Science Foundation NSF IIS-02-09199 and NSF IIS-03-08001.