S.S. Iyengar, *Louisiana State University, and*
N. Balakrishnan, *Indian Institue of Science*

Signal Processing for Mining Information

Data mining is "the data-driven discovery and modeling of hidden patterns in large volumes of data." Data mining differs from retrospective technologies because it produces models that capture and represent hidden patterns in the data. Space images have inspired researchers to study the intrinsic use of information gathering and information delivery. Remote sensing is one of the major applications and can have an impact on the next millennium. Data mining is the discovery of patterns, associations, anomalies, and statistically significant structures in data. It is a multidisciplinary field, borrowing and enhancing ideas from diverse areas such as statistics, signal and image processing, image understanding, mathematical optimization, computer vision, and pattern recognition. It is the process of nontrivial extraction of implicit, previously unknown, and potentially useful information from voluminous data. Advances in various fields including DNA sequencing, bioinfomatics, e-commerce, fraud detection, knowledge management remote sensing images, GIS, digital cartography, stock investment and prediction analysis, and real-time decision making have fueled it. Scientists, especially those working in the signal and image processing areas, generally have now realized that they are ill equipped to handle the overwhelming amounts of data emanating from a host of sources.

With sensors becoming omnipresent, computers simulating complex processes at an unprecedented pace and sensor signal data storage capabilities improving constantly, petabyte-scale data sets are becoming an everyday usage. All this has led to innovative application of data mining techniques to novel and challenging problems. The bottom line is how to identify and exploit structure and patterns in the vast swirl of information that has been collected for the sake of new insights, scientific discovery, and enhancements.

In This Issue

▲ "Secure Sensor Information Management and Mining" by B. Thuraisingham
▲ "Visual Data Mining for Modeling Prior Distributions in Morphometry" by A. Machado, J. Gee, and M. Campos
▲ "A Generic Applied Evolutionary Hybrid Technique" by G. Beligiannis, L. Skarlas, and S. Likothanassis
▲ "Estimating the Gradient Threshold in the Perona-Malik Equation" by F. Voci, S. Eiho, N. Sugimoto, and H. Sekiguchi
▲ "Adjustment of Nonuniform Sampling Locations in Spatial Data Sets" by M. Kamiyama and T. Higuchi
▲ "Kernel Methods and Their Potential Use in Signal Processing" by F. Peréz-Cruz and O. Bousquet
▲ "Mining Information from Event-Related Recordings" by N.A. Laskaris, S.D. Fotopoulos, and A. Ioannides
▲ "Characterization of Protein Secondary Structure" by M.K. Ganapathiraju, J. Klein-Seetharaman, N. Balakrishnan, and R. Reddy

Overview of the Articles

This special section presents some new theoretical results in data mining and knowledge discovery from databases and representative applications in different facets of signal processing. More specifically, we present a discussion of articles in the following categories: secure information management, visual data mining, modeling and information mining, the scale-space theory, spatial data mining, kernel based algorithms in signal processing, data mining framework for intelligent single-trial analysis, and finally characterization of protein secondary structure.

The first article, "Secure Sensor Information Management and Mining," focuses on issues and data structuring challenges for secure information sensor management. This article provides a nice overview of the management of mining sensor data schemes including security for sensor databases on a real-time basis. Paradigms for the tradeoffs between security and real-time processing are discussed in this article.

The second article "Visual Data Mining for Modeling Prior Distributions in Morphometry" presents a novel method for visual

data mining based on exploratory factor analysis. These methods have been applied to an MRI study of the human corpus callosum, and the results are outstanding in terms of revealing differences in the callosal morphology between male and female samples.

The third article, "A Generic Applied Evolutionary Hybrid Technique," presents a generic applied evolutionary hybrid technique that combines the effectiveness of adaptive multimodel partitioning filters and genetic algorithms. The method can be applied to both linear and nonlinear real-world data, and the techniques are very robust in terms of computationally efficiency.

The fourth article, "Estimating the Gradient Threshold in the Perona-Malik Equation," discusses an impressive method of classical scale-space analysis applied to an image segmentation problem. The authors provide a study on the segmentation of microcalcifications on X-ray biomedical images.

The fifth article, "Adjustment of Nonuniform Sampling Locations in Spatial Data Sets," discusses an algorithm for adjusting sampling locations in one spatially discretized data set with those in another when the value differences between these sets are caused by sampling intervals. This algorithm is based on dynamic programming, and this method can be used for applications also.

The sixth article, "Kernel Methods and Their Potential Use in Signal Processing," uses the notion of kernels for many signal processing problems. A nice guideline is provided for future development in this area of kernel methods for signal processing problems.

The seventh article, "Mining Information from Event-Related Recordings," presents a user-friendly data mining framework using the ideas from both pattern analysis and graph theory. This is a very exciting article with some new ideas in the context of signal processing.

The final article, "Characterization of Protein Secondary Structure," presents a very elegant method of characterizing a protein secondary structure using a sequence semantic analysis alone without the use of any evolutionary information or global optimization. This method could be of interest to designers of drugs from pharmaceutical industries.

Acknowledgments

We gratefully acknowledge Area Editor Petar M. Djuric´ and Editor-in-Chief Ray Liu for supporting this special section on signal processing for mining information. We would like to express our sincere appreciation to all those persons who have helped to make this special issue possible, especially to the reviewers for their hard work. This effort is partially funded by the National Science Foundation to Prof. Iyengar.