method could lead to better learning performance and "front-loaded" cost savings. This method is especially pertinent to applications such as fraud detection, where a random initial sample tends to be biased and needs to be obtained at a substantial cost.

3 - Simulating Non-stationary, Non-Poisson, Non-renewal Arrival Processes

Barry L Nelson, Professor, Northwestern University, Department of Industrial Engr. & Mgmt. Sci., 2145 Sheridan Road, C210, Evanston, IL, 60208-3119, United States of America, nelsonb@northwestern.edu, Ira Gerhardt

Every simulation software product includes renewal arrival processes, and many support non-stationary Poisson arrivals. But real arrival processes can be non-stationary, non-Poisson and non-renewal. We present a convenient framework for fitting and simulating arrival processes that may have time-dependent arrival rates, may be more variable or more regular than a Poisson process, and may exhibit dependence among arrivals.

4 - Algorithms for Poisson and Negative Binomial Random Vector Generation

Raghu Pasupathy, Assistant Professor, Virginia Tech, 221 Durham Hall, Blacksburg, VA, 24061, pasupath@vt.edu, Kaeyoung Shin

We present fast algorithms for Poisson random vector generation. These algorithms have complete coverage in two dimensions, and rigorous error control for robust implementation. Furthermore, they have demonstrably faster preprocessing and generation times compared to NORTA. We will also discuss simple extensions for generating negative binomial random vectors through a well-known transformation that uses the gamma distribution.

5 - Multivariate Inputs with Time-varying Distributional Properties

Bahar Biller, Assistant Professor, Carnegie Mellon University, Tepper School of Business, 5000 Forbes Avenue, Posner Hall 360, Pittsburgh, PA, 15213, United States of America, billerb@andrew.cmu.edu, Jim Foster

Although the joint distributional properties of multivariate time-series input processes often vary with time, most of the simulation input models have been developed for stationary processes. We present a model for input processes with possibly many components having time-varying joint distributional properties. This model further generalizes some of the well-known conditional heteroskedastic models used routinely for financial time-series modeling.

■ SA07

C-Room 24B, Upper Level

Data Mining Algorithms for Parsimonious Modeling

Sponsor: Data Mining
Sponsored Session

Chair: Seoung Bum Kim, Assistant Professor, Korea University, Anam-dong Seongbuk-Gu, Seoul, Korea, Republic of, sbkim1@korea.ac.kr

1 - How to Integrate the Diverse Measures for Hospital Fraud Detection?

Hyunjung Shin, Assistant Professor, Ajou University, Suwon, Korea, Republic of, shin@ajou.ac.kr, Junwoo Lee

To detect fraudulent and abusive bill claims of medical care providers, a variety of indexes have developed and evaluated diverse aspects of bill claim pattern. When taking all of indexes into consideration, however, it becomes confusing to find out which index is of more importance than others, and even more difficult if the respective results are significantly discordant. To avoid the ambiguities, we propose a method integrating the diverse degrees of anomaly based on 2007 Korean HIRA data.

2 - A One-class Classification Based Fault Isolation Method for Multivariate Process Diagnosis

Thuntee Sukchotrat, Research Faculty, University of Texas at Arlington, United States of America, thuntee.sukchotrat@mavs.uta.edu, Seoung Bum Kim

A number of fault isolation methods have been developed to identify the contribution of alarms signaled from multivariate control charts. However, most of them require a distributional assumption that restricts their applicability to a wide range of problems. To overcome such limitation, we propose a nonparametric fault isolation approach that decomposes the monitoring statistics obtained from a one-class classification algorithm into components that reflect the contribution of each variable.

3 - Unsupervised Feature Selection Using Weighted Principal Components

Panaya Rattakorn, PhD Candidate, University of Texas at Arlington, Arlington, TX, United States of America, panaya.rattakorn@mavs.uta.edu, Seoung Bum Kim

We proposed an unsupervised feature selection method that combines weighted principal components (WPCs) with thresholding algorithms. Each coefficient of the WPCs represents the importance of each individual feature. To identify the significant coefficients, we proposed two thresholding algorithms: a recursive thresholding algorithm and a bootstrap thresholding algorithm. Our experimental results with simulated data and real datasets demonstrated the effectiveness of the proposed methods.

4 - PCA-Based Feature Selection Method for NDP Approximation to the CRL Scheduling Problem

Jin Young Choi, Assistant Professor, Ajou University, Suwon, Korea, Republic of, choijy@ajou.ac.kr, Seoung Bum Kim

We present a principal component analysis (PCA)-based feature selection method for the neuro-dynamic programming (NDP) approximation to the capacitated re-entrant line scheduling problem. In particular, the PCA-based feature selection method is examined for its potential efficacy of the approximation by generating some example capacitated re-entrant lines.

■ SA08

C-Room 24C, Upper Level

High Dimensional Variable Selection with L1 Type Penalties

Sponsor: Data Mining
Sponsored Session

Chair: Gareth James, Associate Professor, University of Southern California, 522 Hoffman Hall, Information and Operations Management, Los Angeles, CA, 90089, United States of America, gareth@usc.edu

1 - Algorithms for Very Large Scale L1 Minimization and Related Problems

Emmanuel Candes, CalTech, emmanuel@acm.caltech.edu, Jerome Bobin, Stephen Becker

This talk introduces novel algorithms for L1 minimization and other nonsmooth norms such as the nuclear norm. These algorithms are based on ideas from Nesterov, namely, accelerated descent methods and smoothing techniques. We demonstrate the effectiveness of these methods on several examples taken in a variety of different fields.

2 - Forward-LASSO with Adaptive Shrinkage

Peter Radchenko, Assistant Professor, University of Southern California, Los Angeles, CA, Peter.Radchenko@marshall.usc.edu, Gareth James

Both Forward Selection and the Lasso provide computationally feasible methods for performing variable selection in high dimensional regression problems involving many predictors. We propose a new method we call Forward-Lasso Adaptive Shrinkage (FLASH), which incorporates the two approaches as special cases. We provide theoretical justifications and also demonstrate on an extensive set of simulations that FLASH generally outperforms many competing approaches.

3 - A Unified Approach to Model Selection and Sparse Recovery Using Regularized Least Squares

Jinchi Lv, Assistant Professor, University of Southern California, IOM Department, HOH 504, University of Southern California, Los Angeles, CA, 90089, United States of America, jinchilv@marshall.usc.edu, Yingying Fan

In this paper we study the properties of regularization methods in model selection and sparse recovery under the unified framework of regularized least squares (RLS) with concave penalties. For model selection, we establish conditions under which a RLS estimator enjoys the nonasymptotic weak oracle property, where the dimensionality can grow exponentially with sample size. For sparse recovery, we present a sufficient condition that ensures the recoverability of the sparsest solution.

4 - Partial Correlation Estimation by Joint Sparse Regression Models

Ji Zhu, Associate Professor, University of Michigan, 439 West Hall, 1085 South U Ave, Ann Arbor, MI, 48109, United States of America, jizhu@umich.edu

In this talk, we propose a computationally efficient approach for selecting non-zero partial correlations under the high-dimension-low-sample-size setting. This method assumes the overall sparsity of the partial correlation matrix and employs sparse regression techniques for model fitting. It is shown that our method performs well in both non-zero partial correlation selection and the identification of hub variables, and also outperforms two existing methods.