

# Workshop Title: KDD 2008 Workshop on Data Mining using Matrices and Tensors

Workshop Homepage: <http://www.cs.fiu.edu/~taoli/kdd08-workshop/workshop.htm>

## 1. Introduction

The field of pattern recognition, data mining and machine learning increasingly adapt methods and algorithms from advanced matrix computations, graph theory and optimization. Prominent examples are spectral clustering, non-negative matrix factorization, Principal component analysis (PCA) and Singular Value Decomposition (SVD) related clustering and dimension reduction, tensor analysis, L-1 regularization, etc. Compared to probabilistic and information theoretic approaches, matrix-based methods are fast, easy to understand and implement; they are especially suitable for parallel and distributed-memory computers to solve large scale challenging problems such as searching and extracting patterns from the entire Web. Hence the area of data mining using matrices and tensors is a popular and growing area of research activities.

This workshop is a continuation of the theme of [Stanford Workshop on Algorithms for Modern Massive Data Sets](#). This workshop will present recent advances in algorithms and methods using matrix and scientific computing/applied mathematics for modeling and analyzing massive, high-dimensional, and nonlinear-structured data. One main goal of the workshop is to bring together leading researchers on many topic areas (e.g., computer scientists, computational and applied mathematicians) to assess the state-of-the-art, share ideas and form collaborations. We also wish to attract practitioners who seek novel ideas for applications.

## 2. Topic areas

Topic areas for the workshop include (but are not limited to) the following:

### Methods and algorithms:

- Principal Component Analysis and Singular value decomposition for clustering and dimension reduction
- Nonnegative matrix factorization for unsupervised and semi-supervised learning
- Spectral graph clustering
- L-1 Regularization and Sparsification
- Sparse PCA and SVD
- Randomized algorithms for matrix computation
- Web search and ranking algorithms
- Tensor analysis: Rank-1 Decomposition, PARAFAC/CANDECOMP, GLRAM/2DSVD, Tucker decompositions (e.g., the Higher-Order SVD)
- GSVD for classification

- Latent Semantic Indexing and other developments for Information Retrieval
- Linear, quadratic and semi-definite Programming
- Non-linear manifold learning and dimension reduction
- Computational statistics involving matrix computations
- Feature selection and extraction
- Graph-based learning (classification, semi-supervised learning and unsupervised learning)

### **Application areas**

- Information search and extraction from Web
- Text processing and information retrieval
- Image processing and analysis
- Genomics and Bioinformatics
- Scientific computing and computational sciences
- Social Networks

### **3. Paper Submission**

The electronic submission Web site for research papers is available at:  
<http://www.easychair.org/conferences/?conf=dmmt08>.

Papers should be at most 10 pages long, single-spaced, in KDD conference format, in font size 10 or larger with 1-inch margins on all sides.

### **4. Important Dates**

- **June 10, 2008:** Electronic submission of full papers
- **June 17, 2008 :** Author notification
- **June 20, 2008:** Submission of Camera-ready papers
- **August 24, 2008:** Workshop in Las Vegas, USA

### **5. Workshop Organizers**

Chris Ding, University of Texas at Arlington  
Tao Li, Florida International University  
Shenghuo Zhu, NEC Laboratories America

Note: for inquiries please send e-mail to [taoli AT cs.fiu.edu](mailto:taoli AT cs.fiu.edu).