

SDM09 Accepted Papers and Posters

(Alphabetize by the title of the papers and posters)

Title	Authors
Papers	
A Bayesian Approach Toward Finding Communities and Their Evolutions in Dynamic Social Networks	Tianbao Yang, Yun Chi, Shenghuo Zhu, Yihong Gong and Rong Jin
A Family of Large Margin Linear Classifiers and Its Applications to Feature-changing Problems	Jianqiang Shen and Thomas Dietterich
A Framework for Exploring Categorical Data	Varun Chandola, Shyam Boriah and Vipin Kumar
Adaptive Concept Drift Detection	Anton Dries and Ulrich Rückert
AMORI: A Metric-based One Rule Inducer	Niklas Lavesson and Paul Davidsson
An Anyspace Indexing Algorithm with Applications to Sensor Data Mining	Lexiang Ye, Xiaoyue Wang, Eamonn Keogh and Agenor Mafra-Neto
Application of Bayesian Partition Models in Warranty Data Analysis	Markus Mueller, Christoph Schlieder and Axel Blumenstock
Bayesian Cluster Ensembles	Hongjun Wang, Hanhuai Shan and Arindam Banerjee
Constraint-based Subspace Clustering	Elisa Fromont, Adriana Prado and Celine Robardet
Content-based Social Network Exploration Through Multi-way Visual Analysis	Jimeng Sun, Spiros Papadimitriou, Ching-Yung Lin, Nan Cao, Shixia Liu and Weihong Qian
CORE: Nonparametric Clustering of Large Numeric Databases	Andrej Taliun, Michael Böhlen and Arturas Mazeika
DensEst: Density Estimation for Data Mining in High Dimensional Spaces	Emmanuel Müller, Ira Assent, Ralph Krieger, Stephan Günnemann and Thomas Seidl

Detecting Communities in Social Networks using Max-Min Modularity	Jiyang Chen, Osmar Zaiane and Randy Goebel
Discovering Substantial Distinctions among Incremental Bi-Clusters	Faris Alqadah and Raj Bhatnagar
Discretized Spatio-Temporal Scan Window	Seyed H. Mohammadi, Vandana P. Janeja and Aryya Gangopadhyay
Efficient Active Learning with Boosting	Zheng Wang, Yangqiu Song and Changshui Zhang
Efficient Computation of Partial-Support for Mining Interesting Itemsets	Ardian Kristanto Poernomo and Vivekanand Gopalkrishnan
Efficient Discovery of Interesting Patterns Based on Strong Closedness	Mario Boley, Tamas Horvath and Stefan Wrobel
Efficient Multiplicative updates for Support Vector Machines	Vamsi Potluru, Sergey Plis, Morten Morup, Vince Calhoun and Terran Lane
Event Discovery in Time Series	Dan Preston, Pavlos Protopapas and Carla Brodley
Finding links and initiators: a graph reconstruction problem	Heikki Mannila and Evimaria Terzi
FuncICA for Time Series Pattern Discovery	Nishant Mehta and Alexander Gray
GAD: General Activity Detection for Fast Clustering on Large Data	Xin Jin, Sangkyum Kim, Jiawei Han, Liangliang Cao and Zhijun Yin
Grammar Mining	Siegfried Nijssen and Luc De Raedt
Graph Generation with Prescribed Feature Constraints	Xiaowei Ying and Xintao Wu
Hierarchical Linear Discriminant Analysis for Beamforming	Jaegul Choo, Barry L. Drake and Haesun Park

High Performance Parallel/Distributed Biclustering Using Barycenter Heuristic	Waseem Ahmad, Arifa Nisar, Wei-Keng Liao and Alok Choudhary
Hybrid Clustering of Text Mining and Bibliometrics Applied to Journal Sets	Xinhai Liu, Shi Yu, Yves Moreau, Bart De Moor, Wolfgang Glänzel and Frizo Janssens
Identifying Unsafe Routes for Network-Based Trajectory Privacy	Aris Gkoulalas-Divanis, Vassilios Verykios and Mohamed Mokbel
Improved Mean-Shift Clustering via Query Set Compression	Xiaotong Yuan, Ran He and Baogang Hu
Integrated KL (K-means - Laplacian) Clustering: A New Clustering Approach by Combining Attribute Data and Pairwise Relations	Fei Wang, Chris Ding and Tao Li
Learning Random-Walk Kernels for Protein Remote Homology Identification and Motif Discovery	Renqiang Min, Rui Kuang, Anthony Bonner and Zhaolei Zhang
Link Propagation: A Fast Semi-supervised Learning Algorithm for Link Prediction	Hisashi Kashima, Tsuyoshi Kato, Yoshihiro Yamanishi, Masashi Sugiyama and Koji Tsuda
Local Relevance Weighted Maximum Margin Criterion for Text Classification	Quanquan Gu and Jie Zhou
Multi-Modal Hierarchical Dirichlet Process Model for Predicting Image Annotation and Image-Object Label Correspondence	Oksana Yakhnenko and Vasant Honavar
Multi-topic based Query-oriented Summarization	Jie Tang, Limin Yao and Dewei Chen

Near-optimal supervised feature selection among frequent subgraphs	Marisa Thoma, Hong Cheng, Arthur Gretton, Jiawei Han, Hans-Peter Kriegel, Alexander J. Smola, Le Song, Philip S. Yu, Xifeng Yan and Karsten M. Borgwardt
Optimal Distance Bounding on Time-Series Data	Michail Vlachos, Serdar Kozat and Philip S Yu
Outlier Detection with Globally Optimal Exemplar-Based GMM	xingwei yang, longin Jan Latecki and Dragoljub Pokrajac
Parallel Large Scale Feature Selection for Logistic Regression	Sameer Singh, Jeremy Kubica, Scott Larsen and Daria Sorokina
Polynomial-Delay and Polynomial-Space Algorithms for Mining Closed Sequences, Graphs, and Pictures in Accessible Set Systems	Hiroki Arimura and Takeaki Uno
Positive Unlabeled Learning for Data Stream Classification	Xiao-Li Li, Philip S. Yu, Bing Liu and See-Kiong Ng
Prior-Free Rare Category Detection	Jingrui He and Jaime Carbonell
Privacy Preservation in Social Networks with Sensitive Edge Weights	Lian Liu, Jie Wang, Jinze Liu and Jun Zhang
Proximity-Based Anomaly Detection using Sparse Structure Learning	Tsuyoshi Ide, Aurelie Lozano, Naoki Abe and Yan Liu
Scalable Distributed Change Detection from Astronomy Data Streams using Local, Asynchronous Eigen Monitoring Algorithms	Kamalika Das, Kanishka Bhaduri, Sugandha Arora, Wesley Griffin, Kirk Borne, Chris Giannella and Hillol Kargupta
Straightforward Feature Selection for Scalable Latent Semantic Indexing	Jun Yan, Shuicheng Yan, Ning Liu and Zheng Chen
The Metric Dilemma: Competence-Conscious Associative Classification	Adriano Veloso, Mohammed Zaki, Wagner Meira Jr. and Marcos Gonçalves

Time-decayed Correlated Aggregates over Data Streams
Graham Cormode, Srikanta Tirthapura and Bojian Xu

Topic Cube: Topic Modeling for OLAP on Multidimensional Text Databases
Duo Zhang, Chengxiang Zhai and Jiawei Han

Top-K Correlative Graph Mining
Yiping Ke, James Cheng and Jeffrey Xu Yu

Toward Optimal Ordering of Prediction Tasks
Abhimanyu Lad, Yiming Yang and Rayid Ghani

Travel-Time Prediction using Gaussian Process Regression: A Trajectory-Based Approach
Tsuyoshi Ide and Sei Kato

Twin Vector Machines for Online Learning on a Budget
Zhuang Wang and Slobodan Vucetic

Understanding Importance of Collaborations in Co-authorship Networks: A Supportiveness Analysis Approach
Yi Han, Bin Zhou, Jian Pei and Yan Jia

Posters

A Bayesian Approach to Graph Regression with Relevant Subgraph Selection
Silvia Chiappa, Hiroto Saigo and Koji Tsuda

A Hybrid Data Mining Metaheuristic for the p-Median Problem
Alexandre Plastino, Erick Fonseca, Richard Fuchshuber, Simone Martins, Alex Freitas, Martino Luis and Said Salhi

A new constraint for mining sets in sequences
Boris Cule, Bart Goethals and Celine Robardet

A Re-Evaluation of the Over-Searching Phenomenon in Inductive Rule Learning
Frederik Janssen and Johannes Fürnkranz

A Semi-Supervised Framework for Feature Mapping and Multiclass Classification
Bo Chen, Wai Lam, Ivor W Tsang and Tak-lam Wong

Aligned Graph Classification with Regularized Logistic Regression	Brian Quanz and Jun Huan
An Entity Based Model for Coreference Resolution	Michael Wick, Aron Culotta, Khashayar Rohanimanesh and Andrew McCallum
Analyses for Service Interaction Networks with applications to Service Delivery	S Kameshwaran, Sameep Mehta, Vinayaka Pandit, Gyana Parija, Sudhanshu Singh and N. Viswanadham
Change-Point Detection in Time-Series Data by Direct Density-Ratio Estimation	Yoshinobu Kawahara and Masashi Sugiyama
Context Aware Trace Clustering: Towards Improving Process Mining Results	Jagadeesh Chandra Bose R.P and Wil M.P. van der Aalst
Detection and Characterization of Anomalies in Multivariate Time Series	Haibin Cheng, Pang-Ning Tan, Christopher Potter and Steven Klooster
Discovery of Geospatial Discriminating Patterns from Remote Sensing Datasets	Wei Ding, Tomasz Stepinski and Josue Salazar
Diversity-based Weighting Schemes for Clustering Ensembles	Francesco Gullo, Andrea Tagarelli and Sergio Greco
Divide and Conquer Strategies for Effective Information Retrieval	Jie Chen and Yousef Saad
Efficient Secure Computation Protocols for Privacy Preserving Data Mining	Ke Zhai, Wee Keong Ng, Andre Ricardo Herianto and Shuguo Han
Exact Discovery of Time Series Motifs	Abdullah Mueen, Eamonn Keogh, Qiang Zhu, Sydney Cash and Brandon Westover
Exploiting Semantic Constraints for Estimating Supersenses with CRFs	Gerhard Paass and Frank Reichartz

Feature Weighted SVMs Using Receiver Operating Characteristics	Shaoyi Zhang, M. Maruf Hossain, Md. Rafiul Hassan, James Bailey and Kotagiri Ramamohanarao
FEDRA: A Fast and Efficient Dimensionality Reduction Algorithm	Panagis Magdalinos, Christos Doulkeridis and Michalis Vazirgiannis
Finding Representative Rules	Warren Davis, Peter Schwarz and Evimaria Terzi
FutureRank: Ranking Scientific Articles by Predicting their Future PageRank	Hassan Sayyadi and Lise Getoor
Highlighting Diverse Concepts in Documents	Kun Liu, Evimaria Terzi and Tyrone Grandison
Identifying Information-Rich Subspace Trends in High-Dimensional Data	Snehal Pokharkar and Chandan Reddy
Low-Entropy Set Selection	Hannes Heikinheimo, Jilles Vreeken, Arno Siebes and Heikki Mannila
Measuring Discrimination in Socially-Sensitive Decision Records	Dino Pedreschi, Salvatore Ruggieri and Franco Turini
Mining cohesive patterns from graphs with feature vectors	Flavia Moser, Recep Colak, Arash Rafiey and Martin Ester
Mining Complex Spatio-Temporal Sequence Patterns	Florian Verhein
Mining for Surprise Events within Text Streams	Paul Whitney and Dave Engel
Multi-field Correlated Topic Modeling	Konstantin Salomatin, Yiming Yang and Abhimanyu Lad
Multiple Kernel Clustering	Bin Zhao, James Kwok and Changshui Zhang
MUSK: Uniform Sampling of k Maximal Patterns	Mohammad Hasan and Mohammed Zaki
Noise Robust Classification Based On Spread Spectrum	Joern David

Non-Negative Matrix Factorization, Convexity and Isometry	Nikolaos Vasiloglou, Alexander Gray and David Anderson
Non-Parametric Information-Theoretic Measures of One-Dimensional Distribution Functions for Anomaly Detection in Continuous Series	Paolo D'Alberto and Ali Dasdan
On Maximum Coverage in the Streaming Model & Application to Multi-topic Blog-Watch	Barna Saha and Lise Getoor
On Randomness Measures for Social Networks	Xiaowei Ying and Xintao Wu
On Segment Based Stream Modeling and its Applications	Charu Aggarwal
On the Comparison of Relative Clustering Validity Criteria	Lucas Vendramin, Ricardo Campello and Eduardo Hruschka
Parallel Pairwise Clustering	Elad Yom-Tov and Noam Slonim
PICC Counting: Who Needs Joins when you Can Propagate Efficiently?	Jong Wook Kim and K. Selcuk Candan
Providing Privacy through Plausibly Deniable Search	Mummoorthy Murugesan and Chris Clifton
Randomization Techniques for Graphs	Sami Hanhijärvi, Gemma C. Garriga and Kai Puolamäki
Semi-supervised Learning by Sparse Representation	Shuicheng Yan and Huan Wang
ShatterPlots: Fast Tools for Mining Large Graphs	Ana Paula Appel, Deepayan Chakrabarti, Christos Faloutsos, Ravi Kumar, Jure Leskovec and Andrew Tomkins
Spatially Cost-sensitive Active Learning	Alexander Liu, Goo Jun and Joydeep Ghosh
Structure and Dynamics of Research Collaboration in Computer Science	Christian Bird, Earl Barr, Andre Nash, Premkumar Devanbu, Vladimir Filkov and Zhendong Su

Text Categorization with All Substring
Features

Daisuke Okanohara and Jun'ichi Tsujii

The Set Classification Problem and
Solution Methods

Xia Ning and George Karypis

Topic Evolution in a Stream of
Documents

Andrè Gohr, Alexander Hinneburg, Rene
Schult and Myra Spiliopoulou

Tracking user mobility to detect
suspicious behavior

Gaurav Tandon and Philip Chan