SDM07 Accepted Papers and Posters

Title	Authors	Type *
A General Framework for Mining Concept-Drifting Data Streams with Skewed Distributions	Jing Gao, Wei Fan, Jiawei Han and Philip S. Yu	FP
AC-Framework for Privacy-Preserving Collaboration	Wei Jiang and Chris Clifton	FP
Active Learning of Constraints for Semi-supervised Text Clustering	Ruizhang Huang, Wai Lam and Zhigang Zhang	FP
Adaptive Concept Learning through Clustering and Aggregation of Relational Data	Hichem Frigui and Cheul Hwang	FP
An Analysis of Logistic Models: Exponential Family Connections and Online Performance	Arindam Banerjee	FP
Bandits for Taxonomies: A Model-based Approach	Sandeep Pandey, Deepak Agarwal, Deepayan Chakrabarti and Vanja Josifovski	FP
Boosting Optimal Logical Patterns Using Noisy Data	Noam Goldberg and Chung-chieh Shan	FP
Clustering by weighted cuts in directed graphs	Marina Meila and William Pentney	FP
Conical dimension as an intrisic dimension estimator	Xin Yang, Sebastien Michea and Hongyuan Zha	FP
Constraint-Based Pattern Set Mining	Luc De Raedt and Albrecht Zimmermann	FP
Efficient Multiclass Boosting Classification with Active Learning	Jian Huang, Seyda Ertekin, Yang Song, Hongyuan Zha and C. Lee Giles	FP
Fast Counting with AV-Space for Efficient Rule Induction	Linyan Wang and Aijun An	FP

Fast Multilevel Transduction on Graphs	Fei Wang and Changshui Zhang	FP
Fast Newton-type Methods for the Least Squares Nonnegative Matrix Approximation Problem	Dongmin Kim, Suvrit Sra and Inderjit Dhillon	FP
Finding Motifs in Database of Shapes	Xiaopeng Xi, Eamonn Keogh, Li Wei and Agenor Mafra-Neto	FP
Harmonium-Based Models for Semantic Video Representation and Classification	Jun Yang, Yan Liu, Eric Xing and Alexander Hauptmann	FP
Higher Order Orthogonal Iteration of Tensors (HOOI) and its Relation to PCA and GLRAM	Benard Sheehan and Yousef Saad	FP
Identifying Bundles of Product Options using Mutual Information Clustering	Claudia Perlich and Saharon Rosset	FP
Incremental Spectral Clustering With Application to Monitoring of Evolving Blog Communities	Huazhong Ning, Wei Xu, Chi Yun, Yihong Gong and Thomas Huang	FP
Kernel-based Detection of Mislabeled Training Examples	Hamed Valizadegan and Pang-Ning Tan	FP
Lattice based Clustering of Temporal Gene-Expression Matrices	Yang Huang and Martin Farach-Colton	FP
Less is More: Compact Matrix Decomposition for Large Sparse Graphs	Jimeng Sun, Yinglian Xie, Hui Zhang and Christos Faloutsos	FP
Maximizing the Area under the ROC Curve with Decision Lists and Rule Sets	Henrik Bostrom	FP
Maximum Margin Classifiers with Specified False Positive and False Negative Error Rates	Saketha Nath Jagarlapudi and Chiranjib Bhattacharyya	FP
Mining Smooth Evolution of Clusters from Dynamic Data	Yi Wang, Shi-Xia Liu and Li-zhu Zhou	FP
Multi-way Clustering on Relation Graphs	Arindam Banerjee, Sugato Basu and Srujana Merugu	FP

Nonlinear Dimensionality Reduction using Approximate Nearest Neighbors	Erion Plaku and Lydia Kavraki	FP
On Point Sampling Versus Space Sampling for Dimensionality Reduction	Charu Aggarwal	FP
On Privacy-Preservation of Text and Sparse Binary Data with Sketches	Charu Aggarwal and Philip Yu	FP
On Sample Selection Bias and Its Efficient Correction via Model Averaging and Unlabeled Examples	Wei Fan and Ian Davidson	FP
Preventing Information Leaks in Email	Vitor Carvalho and William Cohen	FP
Probabilistic Joint Feature Selection for Multi-task Learning	Tao Xiong, Jinbo Bi, Bharat Rao and Vladimir Cherkassky	FP
RCMap: Efficiently Creating High-Quality Euclidean Embeddings	Arun Qamra and Edward Chang	FP
ROAM: Rule- and Motif-Based Anomaly Detection in Massive Moving Object Data Sets	Xiaolei Li, Jiawei Han, Sangkyum Kim and Hector Gonzalez	FP
Segmentations with rearrangements	Aristides Gionis and Evimaria Terzi	FP
Towards Attack-Resilient Geometric Data Perturbation	Keke Chen and Ling Liu	FP
Semi-Supervised Dimensionality Reduction	Daoqiang Zhang, Zhi-Hua Zhou and Songcan Chen	S Poster Papers
Semi-supervised Feature Selection via Spectral Analysis	Zheng Zhao and Huan Liu	S Poster Papers
Sketching Landscapes of Page Farms	Bin Zhou and Jian Pei	S Poster Papers

Stacked Graphical Models for Efficient Inference in Markov Random Fields	Zhenzhen Kou and William W. Cohen	S Poster Papers
Summarizing Reviews of ``Unequal'' Reviewers	Hady W. Lauw, Ee-Peng Lim and Ke Wang	S Poster Papers
Topic Models over Text Streams: A Study of Batch and Online Unsupervised Learning	Arindam Banerjee and Sugato Basu	S Poster Papers
Understanding and Utilizing the Hierarchy of Abnormal BGP Events	Dejing Dou, Jun Li, Han Qin, Shiwoong Kim and Sheng Zhong	S Poster Papers
Validation of Collaborative Filtering Recommendations in Changing Environments	Olfa Nasraoui, Jeff Cerwinske, Carlos Rojas and Fabio Gonzalez	S Poster Papers
WAT: Finding Top-K Discords in Time Series Database	Yingyi Bu, Tat-wing Leung, Ada Wai-chee Fu, Eamonn Keogh, Jian Pei and Sam Meshkin	S Poster Papers
A Better Alternative to Piecewise Linear Time Series Segmentation	Daniel Lemire	S Poster Papers
A PAC Bound for Approximate Support Vector Machines	Dongwei Cao and Daniel Boley	S Poster Papers
A Practical Point Estimate of False Negatives for Classification Problems where the True Class Label Can Not Be Ascertained for All Instances	Gyorgy Simon and Zhi-Li Zhang	S Poster Papers
A System for Keyword Search on Textual Streams	Vagelis Hristidis, Oscar Valdivia, Michail Vlachos and Philip Yu	S Poster Papers
An incremental data-stream sketch using sparse random projections	Aditya Krishna Menon, Anh Pham, Sanjay Chawla and Anastasios Viglas	S Poster Papers
Approximating Representations for Large Numerical Databases	Szymon Jaroszewicz and Marcin Korzen	S Poster Papers
Are approximation algorithms for consensus clustering worthwhile?	Michael Bertolacci and Anthony Wirth	S Poster Papers
Bursty Feature Representation for Clustering Text Streams	QI HE, Kuiyu Chang, Ee-Peng Lim and Jun Zhang	S Poster Papers

Cascading Behavior in Large Blog Graphs	Jure Leskovec, Mary McGlohon, Christos Faloutsos, Natalie Glance and Matthew Hurst	S Poster Papers
Change-Point Detection using Krylov Subspace Learning	Tsuyoshi Ide and Koji Tsuda	S Poster Papers
Co-Preserving Patterns in Bipartite Partitioning for Topic Identification	Tianming Hu, Jie Zhang, Sam Yuan Sung and Hui Xiong	S Poster Papers
Computing Statistical Profiles o f Active Sites in Proteins	Chang Zhao, Jalal Mahmud, I.V. Ramakrishnan and Subramanyam Swaminathan	S Poster Papers
Discriminating Subsequence Discovery for Sequence Clustering	Jianyong Wang, Yuzhou Zhang, Lizhu Zhou, George Karypis and Charu Aggarwal	S Poster Papers
Distance Preserving Dimension Reduction for Manifold Learning	Hyunsoo Kim, Haesun Park and Hongyuan Zha	S Poster Papers
Distributed Top-k Outlier Detection from Astronomy Catalogs using the DEMAC System	Haimonti Dutta, Chris Giannella, Kirk Borne and Hillol Kargupta	S Poster Papers
Dynamic Algorithm for Graph Clustering Using Minimum Cut Tree	Barna Saha and Pabitra Mitra	S Poster Papers
Fast Best-Match Shape Searching in Rotation Invariant Metric Spaces	Dragomir Yankov, Eamonn Keogh, Li Wei, Xiaopeng Xi and Wendy Hodges	S Poster Papers
Flexible Anonymization For Privacy Preserving Data Publishing: A Systematic Search Based Approach	Bijit Hore, Ravi Chandra Jammalamadaka and Sharad Mehrotra	S Poster Papers
HACS: Heuristic Algorithm for Clustering Subsets	Ding Yuan and Nick Street	S Poster Papers
HP2PC: Scalable Hierarchically-Distributed Peer-to-Peer Clustering	Khaled Hammouda and Mohamed Kamel	S Poster Papers
Learning from Time-Changing Data with Adaptive Windowing	Albert Bifet and Ricard Gavaldà	S Poster Papers
Load Shedding in Classifying Multi-Source Streaming Data: A Bayes Risk Approach	Yijian Bai, Haixun Wang and Carlo Zaniolo	S Poster Papers

Localized Support Vector Machine and Its Efficient Algorithm	Haibin Cheng, Pang-Ning Tan and Rong Jin	S Poster Papers
Mining Visual and Textual Data for Constructing a Multi-Modal Thesaurus	Hichem Frigui and Joshua Caudill	S Poster Papers
On demand Phenotype Ranking through Subspace Clustering	Xiang Zhang, Wei Wang and Jun Huan	S Poster Papers
On Anonymization of String Data	Charu Aggarwal and Philip Yu	S Poster Papers
Poclustering: loseless clustering of dissimilarity data.	Jinze Liu, Qi Zhang, Wei Wang, Leonard McMillan and Jan Prins	S Poster Papers
Rank Aggregation for Similar Items	D. Sculley	S Poster Papers
Robust, Complete, and Efficient Correlation Clustering	Elke Achtert, Christian Böhm, Hans-Peter Kriegel, Peer Kröger and Arthur Zimek	S Poster Papers
Scalable Name Disambiguation using Multi-level Graph Partition	Byung-Won On and Dongwon Lee	S Poster Papers

***FP - Full Papers**
***S Poster Papers - Short Papers / Posters**