



## The 14th Pacific-Asia Conference on Knowledge Discovery and Data Mining

21-24 June, 2010 - Hyderabad, India

---

### Latest NEWS

- PAKDD 2010 in the [news](#)
- 2 Aug 2010: PAKDD photos can be seen here: [Day1](#), [Day2](#), [Day3](#), [Day4](#).
- 2 July 2010: The conference papers are available at the following links.

#### **LNAI 6118:**

Info at: <http://www.springeronline.com/978-3-642-13656-6>

Papers at: <http://www.springerlink.com/content/978-3-642-13656-6/>

#### **LNAI 6119:**

Info at: <http://www.springeronline.com/978-3-642-13671-9>

Papers at: <http://www.springerlink.com/content/978-3-642-13671-9/>

- 15 June 2010: All PAKDD-2010 participants are requested to visit the [general information](#) and [FAQ](#) page.
- 13 June 2010: All PAKDD-2010 participants are requested to submit arrival and departure details [here](#).
- 20 Apr 2010: PAKDD 2010 Program [here](#).
- 7 Apr 2010: Accommodation [details](#) and Venue [details](#) available.
- Proceedings will be published by publication of the proceedings in [Lecture Notes in Artificial Intelligence \(LNAI\)](#).



### About PAKDD

The 14th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2010) is a leading international conference in the areas of data mining and knowledge discovery. It provides an international forum for researchers and industry practitioners to share their new ideas, original research results and practical development experiences from all KDD related areas including data

mining, data warehousing, machine learning, databases, statistics, knowledge acquisition and automatic scientific discovery, data visualization, causal induction and knowledge-based systems. The conference calls for research papers reporting original investigation results and industrial track papers reporting real data mining applications and system development experience. The conference also solicits proposals for tutorials on crucial technologies of knowledge discovery and data mining, and calls for workshop proposals focusing on specific new challenges and emerging issues of knowledge discovery and data mining.

## General Sponsors



DIT/MCIT



CSIR



Yahoo



Persistent



ACE



Trinity College

## Sponsors



AFOSR

**Call for Papers and Proposals**  
**The Fourteenth Pacific-Asia Conference on Knowledge  
Discovery and Data Mining (PAKDD-2010)**

21 – 24 June, 2010, Hyderabad, India  
<http://www.iiit.ac.in/conferences/pakdd2010/>

---

**Aims of the Conference**

The 14th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2010) is a leading international conference in the area of data mining and knowledge discovery. It provides an international forum for researchers and industry practitioners to share their new ideas, original research results and practical development experiences from all KDD related areas including data mining, data warehousing, machine learning, databases, statistics, knowledge acquisition and automatic scientific discovery, data visualization, causal induction and knowledge-based systems.

The conference calls for research papers reporting original investigation results and industrial track papers reporting real data mining applications and system development experience. The conference also solicits proposals for tutorials on crucial technologies of knowledge discovery and data mining, and calls for workshop proposals focusing on specific new challenges and emerging issues of knowledge discovery and data mining.

In previous years, the proceedings of PAKDD have been published in LNCS and we are planning to publish this year proceedings also in LNCS.

**Areas of Interests**

The topics of relevance for the conference papers include but are not limited to the following:

Theoretical Foundations of Data Mining  
Novel Algorithms for Mining  
Association Rules  
Classification and Ranking  
Clustering  
Text Mining  
Machine Learning Methods  
Statistical Methods for Data Mining  
Privacy Preserving Data Mining  
Parallel and Distributed Data Mining  
Interactive and Online Mining  
Graph Mining  
KDD Process and Human Interaction  
Data and Knowledge Visualization  
Knowledge Management  
Mining High Dimensional Data  
Mining Temporal Data  
KDD in Biomedical Domains  
Dynamic Data Mining  
Mining Scientific Databases  
Mining Semi-structured/unstructured Data  
Mining Spatial Data  
Mining Multimedia Data  
Web Data and the Internet  
Mining in Online Gaming and Virtual Worlds  
Social Network Analysis  
Integrated Media Mining  
Security and Intrusion Detection  
Reliability and Robustness Issues  
Mining Noisy Data  
Mining Trends, Opportunities or Risks  
Integration of Data Warehousing, OLAP and Data Mining

Graphic Model Discovery  
Software Warehouse and Software Mining

**Proposals for Tutorials and Workshops:** Tutorial proposals must clearly identify the intended audience. A proposal should include, in English, a description and outline of the proposed contents; names, affiliations, and biographical sketches of the speakers. The intended length of the tutorial (3 or 6 hours) should also be indicated. Submit the proposals by e-mail to Tutorial/Workshop chairs.

---

**Key Dates**

|                           |                  |
|---------------------------|------------------|
| Paper Submission deadline | 30 Nov, 2009     |
| Notification to Authors   | 30 Jan, 2010     |
| Camera-ready copy         | 28 Feb, 2010     |
| Workshop proposals        | 31 Dec, 2009     |
| Tutorial Proposals        | 28 Feb, 2010     |
| Conference Dates          | 21-24 June, 2010 |

---

**Organizing Committee**

**Honorary Chair**

Rajeev Sangal, IIIT, Hyderabad, India

**General Chairs**

Jaideep Srivastava, University of Minnesota, USA  
Masaru Kitsuregawa, University of Tokyo, JAPAN  
P. Krishna Reddy, IIIT, Hyderabad, India

**Program Chairs**

Mohammed J. Zaki, Rensselaer Polytechnic Institute, USA  
Jeffrey Xu Yu, The Chinese University of Hong Kong  
B. Ravindran, IIT Madras, India

**Publication Chair**

Vikram Pudi, IIIT, Hyderabad, India

**Tutorial Chair**

Kamal Karlapalem, IIIT, Hyderabad, India

**Publicity Chair**

Arnab Bhattacharya, IIT Kanpur, India

**Workshop Chair**

Pabitra Mitra, IIT Kharagpur, India

**Local Arrangements Committee**

R K Bagga, IIIT, Hyderabad, India (Chair)  
T Rangunathan, IIIT, Hyderabad, India  
Radhakrishna Pisapati, Infosys SET Labs, Hyderabad, India  
A Govardhan, JNTU, Hyderabad, India  
RBV Subramanyam, Natitonal Insttute of Technology, Warangal

**PAKDD-2010 Secretariat**

International Institute of Information Technology (IIIT),  
Gachibowli, Hyderabad 500 032, Andhra Pradesh, India  
Phone: +91-40-6653 1119; +91-9849750410(Vijay)  
Fax: +91-40-2300 0044; E-mail: pakdd2010@iiit.ac.in  
URL: <http://www.iiit.ac.in/conferences/pakdd2010/>



## Keynote Speakers



**Dr. Wei Ying Ma**

Microsoft Research Asia, China

Homepage: <http://research.microsoft.com/en-us/press/wyma.aspx>

**Title:** Empowering People with Knowledge: The Next Frontier for Web Search

### **Abstract:**

The Web is continuing to evolve at a rapid pace, with the emergence of cloud computing promising to create a new platform for software development and service delivery. One of the greatest opportunities of this new era is the cultivation of a developer ecosystem that can produce millions of micro-vertical services and applications, working together to serve each and every user information need. In this new world, there is an opportunity to build a more powerful and intelligent search engine that both understands what users are trying to accomplish and cooperates with users as they learn, make decisions and take actions. In this talk, I will first discuss some significant trends in cloud computing, before sharing my thoughts on how we can leverage these trends to create both innovative and disruptive technologies for Web search.

### **Biography:**

Dr. Wei-Ying Ma is an Assistant Managing Director at Microsoft Research Asia where he oversees multiple research groups including Web Search and Data Mining, Natural Language Computing, and Human Computer Interaction.

Over the years, under his leadership, Wei-Ying's team of researchers have been recognized as one of the global powerhouses in search, data mining, and multimedia information retrieval related research. The team transferred key technologies into Microsoft's search and online service products. In addition, they published extensively at major conferences such as the SIGIR, WWW, and ACM Multimedia.

Before joining Microsoft in 2001, Wei-Ying was with HP Labs in Palo Alto, California where he worked in the fields of multimedia adaptation and distributed media services infrastructure. From 1994 to 1997, Wei-Ying was engaged in the Alexandria Digital Library project at the University of

California, Santa Barbara. During this time, he developed one of the first web-based image-retrieval systems, Netra, which is regarded as one of the most influential image retrieval systems.

As an active member of the research community, Wei-Ying has published more than 250 papers at international conferences. He currently serves on the editorial boards of ACM Transactions on Information System (TOIS) and ACM/Springer Multimedia Systems Journal. In recent years, he served as program co-chair of WWW 2008, program co-chair of PCM 2007, general co-chair of AIRS 2008, and general co-chair of MMM 2005.

Wei-Ying received a bachelor of science in electrical engineering from the National Tsing Hua University in Taiwan in 1990. He earned a Master of Science degree and doctorate in electrical and computer engineering from the University of California at Santa Barbara in 1994 and 1997, respectively.



**Prof. Y. Narahari**

Indian Institute of Science, Bangalore, India

Homepage: <http://lcm.csa.iisc.ernet.in/hari/>

**Title:** Game Theoretic Approaches to Knowledge Discovery and Data Mining

**Abstract:**

Game theory is replete with brilliant solution concepts such as the Nash equilibrium, the core, the Shapley value, etc. These solution concepts and their extensions are finding widespread use in solving several fundamental problems in knowledge discovery and data mining. The problems include clustering, classification, discovering influential nodes, social network analysis, etc. The

first part of the talk will present the conceptual underpinnings underlying the use of game theoretic techniques in such problem solving. The second part of the talk will delve into two problems where we have recently obtained some interesting results: (a) discovering influential nodes in social networks using the Shapley value and (b) identifying topologies of strategically formed social networks using a game theoretic approach.

**Biography:**

Y. Narahari is currently Professor and Chair at the Department of Computer Science and Automation, Indian Institute of Science, Bangalore. The focus of his current research is to apply Game Theory and Mechanism Design to problems in Internet and Network Economics, Electronic Commerce, and Social Networks. He is the lead author of a research monograph entitled "Game Theoretic Problems in Network Economics and Mechanism Design Solutions" published recently

by Springer, London.

He is a Fellow of IEEE, a J.C. Bose National Fellow, and a Fellow of all leading science and engineering academies in India. The global companies with whom he has collaborated in the recent past include GM R & D, Intel, Infosys Technologies, and Xerox Corporation. He is currently a Senior Editor of the IEEE Transactions on Automation Science and Engineering.



**Prof. Vipin Kumar**

University of Minnesota, Minneapolis, USA.

Homepage: <http://www-users.cs.umn.edu/~kumar/>

**Title:** Discovery of Patterns in Global Earth Science Data using Data Mining

**Abstract:**

The climate and earth sciences have recently undergone a rapid transformation from a data-poor to a data-rich environment. In particular, climate and ecosystem related observations from remote sensors on satellites, as well as outputs of climate or earth system models from large-scale computational platforms, provide terabytes of temporal, spatial and spatio-temporal data.

These massive and information-rich datasets offer huge potential for understanding and predicting the behavior of the Earth's ecosystem and for advancing the science of climate change.

However, mining patterns from Earth Science data is a difficult task due to the spatio-temporal nature of the data. This talk will discuss various challenges involved in analyzing the data, and present some of our work on the design of algorithms for finding spatio-temporal patterns from such data and their applications in discovering interesting relationships among ecological variables from various parts of the Earth. A special focus will be on techniques for land cover change detection (and their use in assessing the impact on carbon cycle) and finding teleconnections between ocean and land variables.

**Biography:**

Vipin Kumar is currently William Norris Professor and Head of Computer Science and Engineering at the University of Minnesota. His research interests include High Performance computing and data mining. He has authored over 250 research articles, and co-edited or coauthored 10 books including the widely used text book "Introduction to Parallel Computing", and "Introduction to Data Mining" both published by Addison-Wesley. Kumar has served as chair/co-chair for over a dozen conferences/workshops in the area of data mining and parallel computing.

In 2001, Kumar co-founded SIAM International Conference on Data Mining and served as its

steering committee chair until 2007. Kumar is a founding co-editor-in-chief of Journal of Statistical Analysis and Data Mining, editor-in-chief of IEEE Intelligent Informatics Bulletin, and series editor of Data Mining and Knowledge Discovery Book Series published by CRC Press/Chapman Hall. Kumar is a Fellow of the AAAS, ACM and IEEE. He received the 2009 Distinguished Alumnus Award from the Computer Science Department, University of Maryland College Park, and 2005 IEEE Computer Society's Technical Achievement Award for contributions to the design and analysis of parallel algorithms, graph-partitioning, and data mining.

---

2010 PAKDD

[Home](#)



## The 14th Pacific-Asia Conference on Knowledge Discovery and Data Mining

21-24 June, 2010 - Hyderabad, India

---

### Accepted Papers

#### Regular Papers

119. [Supervising Latent Topic Model for Maximum-Margin Text Classification and Regression](#)  
Wanhong Xu
  
143. [OddBall: Spotting Anomalies in Weighted Graphs](#)  
Leman Akoglu, Mary McGlohon and Christos Faloutsos
  
509. [Resource-bounded Information Extraction: Acquiring Missing Feature Values On Demand](#)  
Pallika Kanani, Andrew McCallum and Shaohan Hu
  
379. [Efficient Deep Web Crawling Using Reinforcement Learning](#)  
Lu Jiang and Wu Zhaohui
  
19. [A Novel Prototype Reduction Method for the K-Nearest Neighbor Algorithm with  \$K \geq 1\$](#)   
Tao Yang, Longbing Cao and Chengqi Zhang
  
37. [iVAT and aVAT: Enhanced Visual Analysis for Cluster Tendency Assessment](#)  
Liang Wang, Uyen Nguyen, James Bezdek, Christopher Leckie and Kotagiri Ramamohanarao
  
107. [An Efficient GA-Based Algorithm for Mining Negative Sequential Patterns](#)



zhigang zheng, Yanchang Zhao, Ziyue Zuo and Longbing Cao

216. [A Set Correlation Model for Partitional Clustering](#)  
Xuan Vinh Nguyen and Michael E. Houle
276. [An Effective Multi-class ROC](#)  
Md Rafiul Hassan, James Bailey, Kotagiri Ramamohanarao and  
M Maruf Hossain
352. [Ranking sequential patterns with respect to significance](#)  
Robert Gwadera and Fabio Crestani
445. [Generalized Two-Dimensional FLD Method for Face Feature  
Extraction and Recognition](#)  
Shiladitya Chowdhury, Jamuna Kanta Sing, Dipak Kumar Basu  
and Mita Nasipuri
469. [A New Framework for Dissimilarity and Similarity Learning](#)  
Adam Woznica, Alexandros Kalousis and Melanie Hilario
498. [Classification and Novel Class Detection in Data Streams with  
Active Mining](#)  
Mohammad Masud, Jing Gao, Latifur Khan, Jiawei Han and  
Bhavani Thuraisingham
505. [Supervised Learning with Minimal Effort](#)  
Eileen Ni and Charles Ling
71. [Distributed Knowledge Discovery with Non Linear  
Dimensionality Reduction](#)  
Panagis Magdalinos, Michalis Vazirgiannis and Dialecti Valsamou
112. [Hiding Emerging Patterns with Local Recoding Generalization](#)  
Michael Cheng, William Kwok Wai Cheung and Byron Koon Choi
181. [Finding Itemset-Sharing Patterns in a Large Itemset-Associated  
Graph](#)

Mutsumi Fukuzaki, Mio Seki, Hisashi Kashima and Jun Sese

206. [Blog Opinion Retrieval based on Topic-opinion Mixture Model](#)

Peng Jiang, Chunxia Zhang, Qing Yang and Zhendong Niu

224. [EigenSpokes: Surprising Patterns and Scalable Community Chipping in Large Graphs](#)

B. Aditya Prakash, Ashwin Sridharan, Mukund Seshadri, dhar Machiraju and Christos Faloutsos

231. [Valency based Weighted Association Rule Mining](#)

Yun Sing Koh, Russel Pears and Wai Yeap

239. [Orthogonal Nonnegative Matrix Tri-factorization for Semi-supervised Document Co-clustering](#)

Huifang Ma, Weizhong Zhao, Qing Tan and Zhongzhi Shi

269. [Basset: Scalable Gateway Finder in Large Graphs](#)

Hanghang Tong, Spiros Papadimitriou, Christos Faloutsos, Philip Yu and Tina Eliassi-Rad

272. [Semantic-Distance Based Clustering for XML Keyword Search](#)

Weidong Yang and Hao Zhu

289. [Subclass-oriented Dimension Reduction with Constraint Transformation and Manifold Regularization](#)

Bin Tong and Einoshin Suzuki

294. [On Finding the Natural Number of Topics with Latent Dirichlet Allocation: Some Observations](#)

Arun R, Suresh V, Veni Madhavan C E and Narasimha Murty M

370. [Ensemble Learning based on Multi-Task Class Labels](#)

Qing Wang and Liang Zhang

387. [Opinion-Based Imprecise Query Answering](#)

Muhammad Abulaish, Tanvir Ahmad, Jahir Uddin and

Mohammad Najmud Doja

451. [Normalized kernels and similarity indices](#)  
Julien Ah-Pine
468. [Adaptive Matching Based Kernels for Labeled Graphs](#)  
Adam Woznica, Alexandros Kalousis and Melanie Hilario
497. [Rule Synthesizing from Multiple Irrelevant Databases Using Clustering](#)  
Dan He, Xindong Wu and Xingquan Zhu
507. [Probabilistic User Modeling in the Presence of Drifting Concepts](#)  
Vikas Bhardwaj and Ramaswamy Devarajan
24. [Learning Gradients with Gaussian Processes](#)  
Xinwei Jiang, Junbin Gao, Tianjiang Wang and Paul Kwan
233. [Fast Perceptron Decision Tree Learning from Evolving Data Streams](#)  
Albert Bifet, Geoff Holmes, Bernhard Pfahringer and Eibe Frank
246. [Efficiently finding the best parameter for the emerging pattern-based classifier PCL](#)  
thanh son ngo, Mengling Feng, Guimei Liu and Limsoon Wong
302. [Robust Outlier Detection Using Commute Time with Eigenspace Embedding](#)  
Lu Dang Khoa Nguyen and Sanjay Chawla
331. [Mining Antagonistic Communities from Social Networks](#)  
Kuan Zhang, David Lo and Ee-Peng Lim
350. [Subgraph Mining on Directed and Weighted Graphs](#)  
Stephan Gunnemann and Thomas Seidl
397. [A Better Strategy of Discovering Link-Pattern based](#)

[Communities by Classical Clustering Methods](#)

Chen-Yi Lin, Jia-Ling Koh and Arbee L. P. Chen

403. [Ontology-based Mining of Brainwaves: A Sequence Similarity Technique for Mapping Alternative Descriptions of Patterns in Event-Related Potentials \(ERP\) Data](#)

Haishan Liu, Gwen Frishkoff, Robert Frank and Dejing Dou

450. [Correspondence Clustering: An Approach to Cluster Multiple Related Spatial Datasets](#)

Vadeerat Rinsurongkawong and Christoph F. Eick

476. [UNN: A Neural Network for uncertain data classification](#)

Jiaqi Ge and Yuni Xia

519. [Generating Diverse Ensembles to Counter the Problem of Class Imbalance](#)

Thomas Hoens and Nitesh Chawla

**Short Papers**

15. [Feature Subsumption for Sentiment Classification in Multiple Languages](#)

Zhongwu Zhai, Hua Xu, Jun Li and Peifa Jia

31. [Combining Support Vector Machines and the t-Statistic for Gene Selection in DNA Microarray Data Analysis](#)

Tao Yang, Vojislav Kecman, Longbing Cao and Chengqi Zhang

70. [A Robust Seedless Algorithm for Correlation Clustering](#)

Mohammad Aziz and Chandan Reddy

76. [Rough Margin based Core Vector Machines](#)

Gang Niu, Bo Dai, Lin Shang and Yangsheng Ji

83. [Estimate on Expectation for Influence Maximization in Social Networks](#)

Yao Zhang, Qing Gu, Jun Zheng and Daoxu Chen

125. [Mining Wikipedia and Yahoo! Answers for Question Expansion in Opinion QA](#)  
Yajie Miao
152. [Decentralisation of ScoreFinder: A Framework for Credibility Management on User-Generated Contents](#)  
Yang Liao, Aaron Harwood and Ramamohanarao Kotagiri
153. [Computation of Ratios of Secure Summations in Multi-Party Privacy-Preserving Latent Dirichlet Allocation](#)  
Bin Yang and Hiroshi Nakagawa
169. [Compact Margin Machine](#)  
Bo Dai and Gang Niu
191. [Topic Decomposition and Summarization](#)  
Wei Chen, Can Wang, Chun Chen, Lijun Zhang and Jiajun Bu
252. [Classification and Pattern Discovery of Mood in Weblogs](#)  
Thin Nguyen, Dinh Phung, Brett Adams, Truyen Tran and Svetha Venkatesh
330. [GTRACE2: Improving Performance Using Labeled Union Graphs](#)  
Akihiro Inokuchi and Takashi Washio
16. [A Framework for SQL-based Mining of Large Graphs on Relational Databases](#)  
Sriganesh Srihari, Shruti Chandrashekar and Srinivasan Parthasarathy
36. [Anonymizing Transaction Data by Integrating Suppression and Generalization](#)  
Junqiang Liu and Ke Wang
59. [Player Performance Prediction in Massively Multiplayer Online Role-Playing Games \(MMORPGs\)](#)  
Kyong Jin Shim, Richa Sharan and Jaideep Srivastava.
60. [Mining Association Rules in Long Sequences](#)  
Boris Cule and Bart Goethals
99. [Bulk loading Hierarchical Mixture Models for Efficient Stream Classification](#)

Philipp Kranen, Ralph Krieger, Stefan Denker and Thomas Seidl

135. [A New Emerging Pattern Mining Algorithm and its Application in Supervised Classification](#)  
Milton Garcia-Borroto, Jose Francisco Martinez-Trinidad and Jesus Ariel Carrasco-Ochoa
136. [Subseries Join: A Similarity-Based Time Series Match Approach](#)  
Yi Lin
145. [Cost-sensitive Listwise Ranking Approach](#)  
Min Lu, MaoQiang Xie, Yang Wang, Jie Liu and YaLou Huang
150. [Using Association Rules to Solve the Cold-Start Problem in Recommender Systems](#)  
Gavin Shaw, Yue Xu and Shlomo Geva
151. [Hierarchical Clustering of Webpages via Cross-Page and In-Page Link Structures](#)  
Xide Lin
155. [DPSP: Distributed Progressive Sequential Pattern Mining on the Cloud](#)  
Jen-Wei Huang, Su-Chen Lin and Ming-Syan Chen
184. [Classifier Ensemble for Uncertain Data Stream Classification](#)  
Shirui Pan, Kuan Wu, Yang Zhang and Xue Li
278. [Vocabulary Filtering for Term Weighting in Archived Question Search](#)  
Zhao-Yan Ming
286. [Capture of Evidence for Summarization: An application of enhanced Subjective Logic](#)  
Sukanya Manna, B. Sumudu U. Mendis and Tom Gedeon
295. [Satrap: Data and network heterogeneity aware P2P data-mining](#)  
Hock Hee Ang, Vivekanand Gopalkrishnan, Anwitaman Datta, Wee Keong Ng and Chu Hong, Steven Hoi
321. [Online Sampling of High Centrality Individuals in Social Networks](#)

Arun Maiya and Tanya Berger-Wolf

326. [Spatial Clustering with Obstacles Constraints Using Dynamic Piecewise Linear Chaotic Map and Dynamic Nonlinear PSO](#)

XuePing ZHANG

329. [Subsequence Matching of StreamSynopses under the Time Warping Distance](#)

Su-Chen Lin, Mi-Yen Yeh and Ming-Syan Chen

362. [Mining Numbers in Text Using Suffix Arrays and Clustering Based on Dirichlet Process Mixture Models](#)

Minoru Yoshida, Issei Sato, Hiroshi Nakagawa and Akira Terada

398. [SkyDist: Data Mining on Skyline Objects](#)

Christian Bohm, Annahita Oswald, Claudia Plant, Michael Plavinski and Bianca Wackersreuther

467. [Summarizing Multidimensional Data Stream: A Hierarchy-Graph-Based Approach](#)

Yoann Pitarch, Anne Laurent and Pascal Poncelet

473. [Answer Diversification for Complex Question Answering on the Web](#)

Palakorn Achananuparp, Xiaohua Hu and Christopher Yang

483. [Semi-Supervised Tag Recommendation - Using Untagged Resources to Mitigate Coldstart Problems](#)

Christine Preisach, Leandro Balby Marinho and Lars Schmidt-Thieme

484. [TWave: High-Order Analysis of Spatiotemporal Data](#)

Michael Barnathan, Vasileios Megalooikonomou, Christos Faloutsos, Feroze Mohamed and Scott Faro

486. [Relationship Between Diversity and Correlation in Multi-Classifer Systems](#)

Kuo-Wei Hsu and Jaideep Srivastava

540. [Analyzing The Role Of Dimension Arrangement For Data Visualization in Radviz](#)

Luigi Di Caro, Vanessa Frias Martinez and Enrique Frias Martinez

73. [Satisfying Privacy Requirements: One Step Before Anonymization](#)

Xiaoxun Sun, Hua Wang and Jiuyong Li

123. [Mining Closed Episodes from Event Sequences Efficiently](#)

Wenzhi Zhou, Hongyan Liu and Hong Cheng

81. [BoostML: An Adaptive Metric Learning for Nearest Neighbor Classification](#)

Nayyar Abbas Zaidi, David McG Squire and David Suter

103. [Efficient Pattern Mining from Uncertain Data with Sampling](#)

Calin Garboni, Toon Calders and Bart Goethals

118. [Mining Trajectory Corridors Using Frechet Distance and Meshing Grids](#)

Haohan Zhu, Jun Luo, Hang Yin, Xiaotao Zhou, Jianping Fan and F. Benjamin Zhan

138. [An Approach for Fast Hierarchical Agglomerative Clustering using Graphics Processors with CUDA](#)

Arul Shalom S.A., Manoranjan Dash and Minh Tue

171. [Data Transformation for Sum Squared Residue](#)

HYUK CHO and Inderjit Dhillon

178. [Multivariate Equi-width Data Swapping For Private Data Publication](#)

Yidong Li and Hong Shen

192. [Fast Orthogonal Nonnegative Matrix Tri-Factorization for Simultaneous Clustering](#)

Zhao Li, Zhenyu Lu and Xindong Wu

227. [Efficient trade-off between speed processing and accuracy in summarizing data stream](#)

Nesrine Gabsi, Fabrice Clerot and Georges Hebrail

259. [Most Significant Substring Mining Based On Chi-square Measure](#)

Sourav Dutta and Arnab Bhattacharya

363. [Relevant Gene Selection Using Normalized Cut Clustering With Maximal Compression Similarity Measure](#)

Rajni Bala, Ramesh Agarwal and Manju Sardana



401. [Integrative Parameter-free Clustering of Data with Mixed Type Attributes](#)  
Christian Bohm, Sebastian Goebel, Annahita Oswald, Claudia Plant, Michael Plavinski and Bianca Wackersreuther
459. [As Time Goes By: Discovering Eras in Evolving Social Networks](#)  
Michele Berlingerio, Michele Coscia, Fosca Giannotti, Anna Monreale and Dino Pedreschi
480. [A Spotting Anomaly framework](#)  
Cuiping Li, Wenlin He and Hong Chen
481. [Privacy-Preserving Network Aggregation](#)  
Troy Raeder, Marina Blanton, Nitesh Chawla and Keith Frikken
537. [Fast discovery of reliable k-terminal subgraphs](#)  
Melissa Kasari, Hannu Toivonen and Petteri Hintsanen



## The 14th Pacific-Asia Conference on Knowledge Discovery and Data Mining

21-24 June, 2010 - Hyderabad, India

---

### Welcome Message

WELCOME TO PAKDD-2010 Delegates!!!

On behalf of Local Arrangements Committee (LAC), we thank you for your registration for 14th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD-2010) being held at Hyderabad from 21 to 24 June 2010.

Program Committee has finalized entire program for all papers for main conference, workshops and tutorials as given on the conference website. The insertion material which will be given as part of your registration kit is free of cost. We have published limited number of copies of PAKDD-2010 Proceedings which will be available to delegates during registration for those only entitled. If few others want to purchase the same, a nominal cost of Rs.5000 can be paid at the registration desk to get a copy.

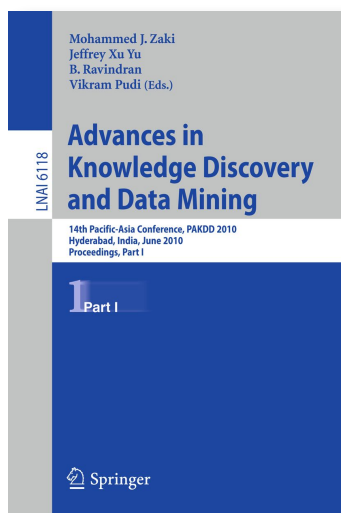
We have planned both reception and banquet dinners at the conference venue keeping in view of rainy season. You are also recommended to carry a rain jacket and a small umbrella, if possible. Do not hesitate to get in touch with PAKDD-2010 Secretariat, if you have any queries. Please go through the [general information](#) for the participants and the [FAQs](#) before contacting us.

P Krishna Reddy

R K BAGGA

General Chair

Chair, Local Arrangements Committee



**M.J. Zaki**, Rensselaer Polytechnic Institute, Troy, NY, USA; **J.X. Yu**, The Chinese University of Hong Kong, China; **B. Ravindran**, IIT Madras, Chennai, India; **V. Pudi**, IIIT Hyderabad, India (Eds.)

## **Advances in Knowledge Discovery and Data Mining, Part I**

14th Pacific-Asia Conference, PAKDD 2010, Hyderabad, India, June 21-24, 2010, Proceedings

- ▶ **up-to-date results**
- ▶ **fast track conference proceedings**
- ▶ **state-of-the-art report**

This book constitutes the proceedings of the 14th Pacific-Asia Conference, PAKDD 2010, held in Hyderabad, India, in June 2010.

1st Edition., 2010, 506 p. 167 illus.

### **Printed book**

#### **Softcover**

- ▶ **70,00 € | £63.99 | \$107.00**
- ▶ **\*74,90 € (D) | 77,00 € (A) | CHF 100.50**

### **eBook**

Available from libraries offering Springer's eBook Collection, or for individual purchase via online bookstores.

**A free preview is available on SpringerLink.**

- ▶ [springer.com/ebooks](http://springer.com/ebooks)

### **MyCopy**

Printed eBook exclusively available to patrons whose library offers Springer's eBook Collection.\*\*\*

- ▶ **€ | \$ 24.95**
- ▶ [springer.com/mycopy](http://springer.com/mycopy)



Order online at [springer.com](http://springer.com) ▶ or for the Americas call (toll free) 1-800-SPRINGER ▶ or email us at: [orders-ny@springer.com](mailto:orders-ny@springer.com). ▶ For outside the Americas call +49 (0) 6221-345-4301 ▶ or email us at: [orders-hd-individuals@springer.com](mailto:orders-hd-individuals@springer.com).

The first € price and the £ and \$ price are net prices, subject to local VAT. Prices indicated with \* include VAT for books; the €(D) includes 7% for Germany, the €(A) includes 10% for Austria. Prices indicated with \*\* include VAT for electronic products; 19% for Germany, 20% for Austria. All prices exclusive of carriage charges. Prices and other details are subject to change without notice. All errors and omissions excepted.

\*\*\* Regional restrictions apply.