2010 12th International Asia-Pacific Web Conference

(APWEB 2010)

Busan, South Korea
6-8 April 2010
# 2010 12th International Asia-Pacific Web Conference

## APWeb 2010

### Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Message from the General Chairs and Program Chairs</td>
<td>xi</td>
</tr>
<tr>
<td>Conference Organizers</td>
<td>xii</td>
</tr>
<tr>
<td>Program Committee Members</td>
<td>xiv</td>
</tr>
<tr>
<td>External Reviewers</td>
<td>xvii</td>
</tr>
<tr>
<td>Message from the USDM Workshop Chairs</td>
<td>xviii</td>
</tr>
</tbody>
</table>

### Keynote Speakers

- **The Negative Byproducts of the Internet**
  - *Won Kim*
  - Page 3

- **Structured Data on the Web**
  - *Alon Y. Halevy*
  - Page 7

### Research Sessions

### Full Papers

#### Full Paper Session 1: Spatial & Temporal Database

- **Aggregate Nearest Keyword Search in Spatial Databases**
  - *Zhicheng Li, Hu Xu, Yansheng Lu, and Ailing Qian*
  - Page 15

- **StickViz: A New Visualization Tool for Phenomenon-Based k-Neighbors Searches in Geosocial Networking Services**
  - *Kyoung-Sook Kim, Koji Zettsu, Yutaka Kidawara, and Yasushi Kiyoki*
  - Page 22

- **Moving K-Nearest Neighbor Query over Obstructed Regions**
  - *Chuanwen Li, Yu Gu, Fangfang Li, and Mo Chen*
  - Page 29

#### Full Paper Session 2: Access Methods & Storage Management

- **A Hybrid Flash Memory SSD Scheme for Enterprise Database Applications**
  - *Byung-Woo Nam, Gap-Joo Na, and Sang-Won Lee*
  - Page 39

- **HashTree: A New Hybrid Index for Flash Disks**
  - *Kai Cui, Peiquan Jin, and Lihua Yue*
  - Page 45
Using Index in the MapReduce Framework .................................................................52
   Mingyuan An, Yang Wang, and Weiping Wang

Full Paper Session 3: Web Search

Suggesting Topic-Based Query Terms as You Type ....................................................61
   Ju Fan, Hao Wu, Guoliang Li, and Lizhu Zhou
An Effective and High-quality Query Relaxation Solution on the Deep Web ......................68
   Jing Shan, Derong Shen, Tiezheng Nie, Yue Kou, and Ge Yu
Applying Taxonomic Knowledge and Semantic Collaborative Filtering to Personalized
Search: A Bayesian Belief Network Based Approach .....................................................75
   Jae-won Lee, Han-joon Kim, and Sang-goo Lee

Full Paper Session 4: Web Service & Workflow

Service Collaboration Network: A Novel Mechanism for Web Service Management ..............85
   Guang Ji, Jie Zhang, Zhuofeng Zhao, and Yanbo Han
Retrieving Cultural Heritage Information with Google Earth ...........................................92
   Tobias Seppenhauser, Alexander Stenzer, and Burkhard Freitag
Dynamically Analyzing Time Constraints in Workflow Systems with Fixed-Date
Constraint .......................................................................................................................99
   Rui Han, Yingbo Liu, Lijie Wen, and Jianmin Wang

Full Paper Session 5: Recommendation

Improving a News Recommendation System in Adapting to Interests of a User
with Storage of a Constant Size ..................................................................................109
   Akito Nishitarumizu, Tsuyoshi Itokawa, Teruaki Kitasuka, and Masayoshi Aritsugi
Graph-Based Recommendation on Social Networks .....................................................116
   Ziqi Wang, Yuwei Tan, and Ming Zhang
Recommendation for Movies and Stars Using YAGO and IMDB .................................123
   Yajie Hu, Ziqi Wang, Wei Wu, Jianzhong Guo, and Ming Zhang

Full Paper Session 6: Clustering & Pattern Mining

Effective Clustering of Dense and Concentrated Online Communities ..........................133
   Phan Nhat Hai and Hyoseop Shin
Multi-View Clustering with Web and Linguistic Features for Relation Extraction ...............140
   Yulan Yan, Haibo Li, Yutaka Matsuo, Zhenglu Yang, and Mitsuru Ishizuka
An Efficient Method for Incremental Mining of Share-Frequent Patterns .......................147
   Chowdhury Farhan Ahmed, Syed Khairuzzaman Tanbeer, and Byeong-Soo Jeong

Full Paper Session 7: Similarity Measures

Local Methods for Estimating SimRank Score ...........................................................157
   Xu Jia, Hongyan Liu, Li Zou, Jun He, Xiaoyong Du, and Yuanzhe Cai
A Space and Time Efficient Algorithm for SimRank Computation ....................................164
   Weiren Yu, Xuemin Lin, and Jiajin Le
An Asymmetric Similarity Measure for Tag Clustering on Flickr

Xiaochen Huang and Ying Zhou

Full Paper Session 8: Similarity Search

Two Novel Adaptive Symbolic Representations for Similarity Search in Time Series Databases

Ninh D. Pham, Quang Loc Le, and Tran Khanh Dang

An Incremental Prefix Filtering Approach for the All Pairs Similarity Search Problem

Hoang Thanh Lam, Dinh Viet Dung, Raffaele Perego, and Fabrizio Silvestri

Efficient Peer-to-Peer Similarity Query Processing for High-dimensional Data

Ye Yuan, Guoren Wang, and Yongjiao Sun

Full Paper Session 9: Query Processing & Optimization 1

Recording How-Provenance on Probabilistic Databases

Ming Gao, Xiangnan He, Cheqing Jin, Xiaoling Wang, and Aoying Zhou

Query-Aware Complex Object Buffer Management in XML Information Retrieval

Qiushi Li, Qiuyue Wang, and Shan Wang

Efficient Common Items Extraction from Multiple Sorted Lists

Wei Lu, Chuitian Rong, Jinchuan Chen, Xiaoyong Du, Gabriel Pui Cheong Fung, and Xiaofang Zhou

Full Paper Session 10: Web Mining 1

Combining the Missing Link: An Incremental Topic Model of Document Content and Hyperlink

Huifang Ma, Zhixin Li, and Zhongzhi Shi

Crawling Online Social Graphs

Shaozhi Ye, Juan Lang, and Felix Wu

Learning to Find Interesting Connections in Wikipedia

Marek Ciglan, Étienne Rivière, and Kjetil Nørvåg

Full Paper Session 11: Multimedia Information Systems

Object Recognition via Adaptive Multi-level Feature Integration

Mei Wang, Yanling Wu, Guangda Li, and Xiangdong Zhou

Efficient and Continuous Near-duplicate Video Detection

Qing Xie, Zi Huang, Heng Tao Shen, Xiaofang Zhou, and Chaoyi Pang

Text-Aided Image Classification: Using Labeled Text from Web to Help Image Classification

Yuan Lin, Yuqiang Chen, Guirong Xue, and Yong Yu

Full Paper Session 12: Query Processing & Optimization 2

Efficient SLCA Computing Service for XML Filtering in IaaS

Chenjing Zhang, Xiaoling Wang, and Aoying Zhou

Computing Large Skylines over Few Dimensions: The Curse of Anti-correlation

Henning Köhler and Jing Yang
A Query Processing Approach Based on Queuing Model for Cyber-Physical Systems ...........................................291
   Fangfang Li, Dejun Yue, and Chuanwen Li

Full Paper Session 13: Web Mining 2
Using Temporal Information to Improve Predictive Accuracy of Collaborative
Filtering Algorithms ...................................................................................................................................................301
   Jia Rongfei, Jin Maozhong, and Liu Chao
Named Entity Resolution in Chinese News Comments on the Web .................................................................307
   Liang Zong, Xiaojun Wan, Lihong Zhao, Jianwu Yang, and Yuqian Wu
ECON: An Approach to Extract Content from Web News Page .......................................................................314
   Yan Guo, Huifeng Tang, Linhai Song, Yu Wang, and Guodong Ding

Full Paper Session 14: Multidimensional Database & OLAP
ParaCube: A Scalable OLAP Model Based on Distributed Aggregate Computing
with Sibling Cubes ..................................................................................................................................................323
   Yansong Zhang, Shan Wang, and Wei Huang
Partitioned Dimension: Modeling the Numerical Dimension in Data Warehouse ........................................330
   Jie Song and Yubin Bao

Short Papers

Short Paper Session 1: New Applications
OntLMS: An Ontology-Based Lifelog Management System ........................................................................341
   Inchul Song, Deokmin Haam, Hangkyu Kim, and Myoung Ho Kim
Entity-Event Lifelog Ontology Model (EELOM) for LifeLog Ontology Schema
Definition .........................................................................................................................................................344
   Sangkeun Lee, Gihyun Gong, and Sang-goo Lee
Classification-Based Skin Aging Analysis Scheme ......................................................................................347
   Young-Hwan Choi, Kyungrok Kim, and Eenjun Hwang
Recommended or Not Recommended? Review Classification through Opinion
Extraction ..................................................................................................................................................350
   Sheng Feng, Ming Zhang, Yanxing Zhang, and Zhihong Deng

Short Paper Session 2: Query Processing & Optimization
Multiple Continuous Queries Optimization for Processing RFID Streaming Data ......................................355
   Haipeng Zhang, Wooseok Ryu, Yongxu Piao, and Bonghee Hong
Skyline Minimum Vector .................................................................................................................................358
   Su Min Jang, Choon Seo Park, and Jae Soo Yoo
Complex Event Detection in Probabilistic Stream ....................................................................................361
   Xu Chuanfei, Lin Shukuan, Wang Lei, and Qiao Jianzhong
Adaptive Top-k Algorithm in SLCA-Based XML Keyword Search ............................................................364
   Hang Yu, Zhihong Deng, Yongqing Xiang, Ning Gao, Ming Zhang, and Shiwei Tang
Short Paper Session 3: Data Mining

MMIR: Mining Multi-scale Intervention Rules in Sub-Complex System ................................................................. 369  
Zheng Jiaoling, Tang Changjie, Qiao Shaojie, Yang Ning, Wang Yue, Chen Yu, and Zhu Jun

Detect Events on Noisy Textual Datasets ................................................................................................................... 372  
Sen Yang, Xueqi Cheng, You Chen, Gaolin Fang, Jin Zhang, and Hongbo Xu

Mining Regular Patterns in Incremental Transactional Databases ............................................................................. 375  
Syed Khairuzzaman Tanbeer, Chowdhury Farhan Ahmed, and Byeong-Soo Jeong

Host-IP Clustering Technique for Deep Web Characterization .................................................................................. 378  
Denis Shestakov and Tapio Salakoski

Short Paper Session 4: Potpourri

Implementation of Autonomous Pricing Mechanism in Grid Computing ................................................................. 383  
S.M. Aminul Haque, Saadat M. Alhashmi, and Rajendran Parthiban

A Cluster-Based Group Key Management Scheme for Wireless Sensor Networks ......................................................... 386  
Yuan Zhang, Yongluo Shen, and SangKeun Lee

SOARWare: A Service Oriented Software Production and Running Environment ....................................................... 389  
Hailong Sun, Xudong Liu, Xiang Li, Jin Zeng, and Zicheng Huang

A Metric for Measuring Members’ Contribution to Information Propagation in Social Network Sites .......................................................... 392  
Jiajia Zheng, Wei Chen, Lijun Zhang, Jiajun Bu, and Chun Chen

Tutorials

Optimizing Query Processing for the Hidden Web ........................................................................................................ 397  
Andrea Calì and Davide Martinenghi

Aggregating, Operating, Sharing and Utilizing Internet-Based Services with the VINCA Approach .......................................................... 398  
Yanbo Han and Zhuofeng Zhao

Industrial Paper

Building Web Page Logical Structure Model towards Effective Metadata Extraction ......................................................... 401  
Baoyao Zhou and Ming Zhang

Workshop Papers

Workshop 1: 3DLife: Bringing the Media Internet to Life

A Multi-Pattern Search Algorithm for Block Motion Estimation in Video Coding ............................................................. 407  
Muhammad Akram and Ebroul Izquierdo

Continuous-tone Watermark Hiding in Halftone Images ................................................................................................ 411  
Jing-Ming Guo and Yun-Fu Liu

Multi-Factor Authentication Using Fingerprints and User-Specific Random Projection ............................................. 415  
Esla Timothy Anzaku, Hosik Sohn, and Yong Man Ro

A Fast Implementation of Arithmetic Coding ........................................................................................................... 419  
Hyoung Joong Kim
Workshop 2: Unstructured Data Management

Session 1: Architecture

laUD-MS: An Extensible System for Unstructured Data Management .........................................................435
   Yuqing Zhu, Naiqiao Du, Hao Tian, and Jianmin Wang
Managing a Large Shared Bank of Unstructured Data by Using Free-Table ..........................................................441
   Xiao Zhang, Xiaoyong Du, Jinchuan Chen, and Shan Wang
A Three Layer System Architecture for Web-Based Unstructured Data Management ...........................................447
   Zhiyong Peng, Hui Wang, Yuwei Peng, Bo Xu, and Zeqian Huang

Session 2: New Applications

Towards Video Management over Relational Database ..................................................................................453
   Hui Li, Xiao Zhang, Shan Wang, and Xiaoyong Du
Phosphor: A Cloud Based DRM Scheme with Sim Card ..................................................................................459
   Peng Zou, Chaokun Wang, Zhang Liu, and Dalei Bao
Domain-oriented Deep Web Data Sources’ Discovery and Identification .........................................................464
   Yingjun Li, Tiezheng Nie, Derong Shen, and Ge Yu
A Heuristic Method for Unstructured Pattern Management over Data Streams .............................................468
   Miao Gaoshan, Li Hongyan, and Wang Tengjiao

Author Index ................................................................................................................................................472