Ford Lumban Gaol (Ed.)

## Recent Progress in Data Engineering and Internet Technology

Volume 1



Editor
Ford Lumban Gaol
Bina Nusantara University
Perumahan Menteng
Jakarta
Indonesia

ISSN 1876-1100 e-ISSN 1876-1119 ISBN 978-3-642-28806-7 e-ISBN 978-3-642-28807-4 DOI 10.1007/978-3-642-28807-4 Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012933326

## © Springer-Verlag Berlin Heidelberg 2013

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

## **Foreword**

In the recent years we are being faced with flooding of data whose handling has promoted rapid advancements in internet technology. This book is a collection of selected papers, all of which were presented at the International Conference on Data Engineering and Internet Technology (DEIT 2011).

This conference, which took place in Bali Dynasty Resort, Bali, Indonesia on 15th–17th March 2011, brought together researchers and scientists from academia, industry, and government laboratories to share their results and identify future research directions in data engineering and internet technology. Topics of interest include, among others: computational algorithms and tools, database management and database technologies, intelligent information systems, data engineering applications, internet security, internet data management, web search, data grids, cloud computing, as well as web-based application.

The book makes fascinating reading and will be not only of great interest to researchers involved in all aspects of Data engineering and Internet Technology, but also it will surely attract more practitioners and academics towards these research areas.

Dr. Ford Lumban Gaol

## **Contents**

Wu Di, Hou Jiuyang, Li Qiang	1
Autonomous and Pervasive Computing-Based Knowledge Service	9
Short-Term Traffic Flow Forecasting Based on Periodicity Similarity Characteristics	15
An Eigenvector-Based Kernel Clustering Approach to Detecting Communities in Complex Networks	23
Distributed Gaussian Mixture Learning Based on Variational Approximations Behrouz Safarinejadian	29
The Function and Relationship of Verbal and Nonverbal Cues in IM Interpersonal Communication	35
Using Version Control System to Construct Ownership Architecture  Documentations	41
DF-ReaL2Boost: A Hybrid Decision Forest with Real L2Boosted Decision Stumps  Zaman Md. Faisal, Sumi S. Monira, Hideo Hirose	47
Magazine Image Retrieval with Camera-Phone	55

VIII Contents

Statistical Clustering and Times Series Analysis for Bridge Monitoring  Data	61
Man Nguyen, Tan Tran, Doan Phan	
Towards Smart Advisor's Framework Based on Multi Agent Systems and Data Mining Methods	73
Automatic LSA-Based Retrieval of Synonyms (for Search Space Extension)	79
Research on Method of Wavelet Function Selection to Vibration Signal Filtering	87
Time Series Subsequence Matching Based on Middle Points and Clipping	93
Effects of Spatial Scale in Cellular Automata Model for Land Use Change	101
Prediction Model Based on PCA - DRKM–RBF	107
A Topic Detection and Tracking System with TF-Density	115
Community Identification of Financial Market Based on Affinity  Propagation	121
Research on Optimization of Target Oil Wells for CO <sub>2</sub> Huff and Puff in Yushulin Oil Well	129
A Secret Embedding Scheme by Means of Re-indexing VQ Codebook upon Image Processing	135
Knowledge Management System: Combination of Experts' Knowledge and Automatic Improvement	143
Evaluating Recommender System Using Multiagent-Based Simulator: Case Study of Collaborative Filtering Simulation	155

On the Constraint Satisfaction Method for University Personal Course Scheduling	163
Factors Affecting Productivity of Fractured Horizontal Wells Li Tiejun, Guo Dali, Tang Zhihao, Ke Xijun	175
Informal Lightweight Knowledge Extraction from Documents Francesco Colace, Massimo De Santo, Paolo Napoletano	181
A Classification Model Using Emerging Patterns Incorporating Item Taxonomy	187
Synchronised Data Logging, Processing and Visualisation Framework for Heterogeneous Sensor Networks	193
Design of SENIOR: A Case Study Using MoGap  Per Karlström, Wenbiao Zhou, Dake Liu	199
GWDL: A Graphical Workflow Definition Language for Business Workflows	205
A Feature Representation and Extraction Method for Malicious Code Detection Based on LZW Compression Algorithm	211
CASM: Coherent Automated Schema Matcher	219
Research on Constructing 3D Geological Model of the Construction Layers in Daxing New City Area of Beijing City	225
Preliminary Study of HGML-Based Virtual Scene Construction and Interaction Display	231
The Next-Generation Search Engine: Challenges and Key Technologies	239
<b>An Efficient k-Anonymization Algorithm with Low Information Loss</b> <i>Md. Nurul Huda, Shigeki Yamada, Noboru Sonehara</i>	249

X Contents

Design and Realization of Data Gateway for Large-Scale Alternately Supervisory System	255
Research on Mass Geospatial Raster Data Processing Based on Map/Reduce Model	261
A Parallel-Enabled Frequency Offset Estimation Scheme for Optical Coherent Receivers	269
Process-Driven Integrated Product Design Platform	275
Segmentation of Audio Visual Malay Digit Utterances Using Endpoint  Detection	281
Data Mining in Clinical Decision Support Systems	287
Spatial Influence Index Effects in the Analysis of Epidemic Data	295
EPON's Research on Transmission of Sampling Value in Process Level of Digital Substation	307
Intelligent Clinical Information Systems for the Cardiovascular Diseases Nan-Chen Hsieh, Huan-Chao Keh, Chung-Yi Chang, Chien-Hui Chan	315
Forecast the Foreign Exchange Rate between Rupiah and US Dollar by Applying Grey Method	321
Simulating Patent Knowledge Contexts	329
Apply Data Mining on Location-Based Healthcare System	337
Using Fuzzy Logic in Virus-Mediated Gene Therapy	343

Towards Using Cached Data Mining for Large Scale Recommender  Systems	349
Swapneel Sheth, Gail Kaiser  BAX-SET PLUS: A Taxonomic Navigation Model to Categorize,  Search and Retrieve Semantic Web Services	359
<b>Text-Transformed Image Classification Based on Data Compression</b> <i>Nuo Zhang, Toshinori Watanabe</i>	365
Numerical Analysis of Touch Mode Capacitive Pressure Sensor Using Graphical User Interface	371
Multithreading Embedded Multimedia Application for Vehicle Blackbox	379
An EMD-Oriented Hiding Scheme for Reversible Data Embedding in Images	385
Gait Control for Guide Dog Robot to Walk and Climb a Step	393
Artificial Neural Network-Based Lot Number Recognition for Cadastral Map	401
The Novel Virtual Reality Fixture Design and Assembly System (VFDAS)	407
Multi-phenomenon Oriented Embedded Software Fault Diagnosis	413
Asymptotic Behavior of Random Walks with Resting State in Ergodic Environments	421
Research on the Spatio-temporal Pass Index Model in Emergency Evacuation	429

XII Contents

On Issues of Multi-path Routing in Overlay-Networks Using Optimization Algorithms	435
Reviews on Intelligent Building Systems in China	441
XBeGene: Scalable XML Documents Generator by Example Based on Real Data	449
A Preliminary Activity Recognition of WSN Data on Ubiquitous Health Care for Physical Therapy	461
<b>Development and Evaluation of Question Templates for Text Mining</b> <i>Norio Ishii, Yuri Suzuki, Takashi Fujii, Hironobu Fujiyoshi</i>	469
Segmentation of Fibro-Glandular Discs in Digital Mammograms Using Log-Normal Distribution	475
An OOAD Model of Program Understanding System's Parser	481
The Unified Modeling Language Model of an Object Oriented  Debugger System	489
A FPGA-Based Real Time QRS Complex Detection System Using Adaptive Lifting Scheme	497
A Novel Temporal-Spatial Color Descriptor Representation for Video Analysis	505
Author Index	511