2013 IEEE 14th International Conference on Information Reuse and Integration

(IRI 2013)

San Francisco, California, USA
14-16 August 2013
TABLE OF CONTENTS

Forward ................................................................................................................................................. xii
Message from Program Co-chairs ....................................................................................................... xiii
Conference Organizers ........................................................................................................................ xv
Program Committee .............................................................................................................................. xvii
Keynote Speech (I) ................................................................................................................................ xx
Keynote Speech (II) ............................................................................................................................. xxiii
Keynote Speech (III) ........................................................................................................................... xxv
Panel .................................................................................................................................................. xxvii
EM-RITE 2013 Program Committee ............................................................................................... xxviii
IRI-HI 2013 Program Committee ................................................................................................... xxix
FMi 2013 Program Committee ......................................................................................................... xxx
DIM 2013 Program Committee ......................................................................................................... xxxii
Technical Papers .................................................................................................................................. xxxiii

Session A11: Information Security & Privacy

FIEP: An Initial Design of A Firewall Information Exchange Protocol ................................................. 1
  Sandeep Reddy Pedditi, Du Zhang and Chung-E Wang
  California State University, Sacramento, USA

Predicting Susceptibility to Social Bots on Twitter ........................................................................... 6
  Randall Wald (1), Taghi M. Khoshgoftaar (1), Amri Napolitano (1) and Chris Sumner (2)
  (1) Florida Atlantic University, USA
  (2) Online Privacy Foundation, USA

Simulation-Based Validation for Smart Grid Environments ................................................................. 14
  Wonkyu Han, Michael Mabey and Gail-Joon Ahn
  Arizona State University, USA

A Moving Target Defense Approach for Protecting Resource-Constrained Distributed Devices ........ 22
  Valentina Casola (1), Alessandra De Benedictis (1) and Massimiliano Albanese (2)
  (1) University of Naples Federico II, Italy
  (2) George Mason University, USA

Session A12: Recommender Systems

Tag-based Top-N Recommendation Using a Pairwise Topic Model ...................................................... 30
Zhengyang Li and Congfu Xu
Zhejiang University, China
Bayesian Model for a Multicriteria Recommender System with Support Vector Regression
Pannawit Samatthiyadikun (1), Atsuhiro Takasu (1, 2) and Saranya Maneeroj (3)
(1) SOKENDAI, Japan
(2) NII, Japan
(3) Chulalongkorn University, Thailand
WS-Rec: Web Services Recommendations using Heterogeneous Criteria
Anna Squicciarini (1), Barbara Carminati (2) and Sushama Karumanchi (1)
(1) Pennsylvania State University, USA
(2) University of Insubria, Varese
Constrained by Limited Information and Recommendation Opportunities: An Exploration and
Exploitation Problem for Recommender Systems
Simon Chan
University College London, UK
(EM-RITE)
Rank Correlation Analysis of NTCIR-10 RITE-2 Chinese Datasets and Evaluation Metrics
Chuan-Jie Lin (1), Cheng-Wei Lee (2), Cheng-Wei Shih (2) and Wen-Lian Hsu (2)
(1) National Taiwan Ocean University, Taiwan
(2) Academia Sinica, Taiwan
Chinese Textual Entailment with Wordnet Semantic and Dependency Syntactic Analysis
Chun Tu and Min-Yuh Day
Tamkang University, Taiwan
Entailment Analysis for Improving Chinese Textual Entailment System
Shih-Hung Wu (1), Shan-Shun Yang (1), Liang-Pu Chen (2), Hung-Sheng Chiu (2) and
Ren-Dar Yang (2)
(1) Chaoyang University of Technology, Taiwan
(2) Institute for Information Industry, Taiwan
Interest Analysis using Social Interaction Content with Sentiments
Lun-Wei Ku and Chung-Chi Huang
Academia Sinica, Taiwan
Clustering and Summarization Topics of Subject Knowledge Through Analyzing Internal Links
of Wikipedia
I-Chin Wu, Chi-Hong Tsai and Yu-Hsuan Lin
Fu-Jen Catholic University, Taiwan
Session A21: Data Mining I
Solving Minimum Cut Based Multi-Label Classification Problem with Semi-Definite
Programming Method
Guangzhi Qu
Oakland University, USA
The Importance of Performance Metrics Within Wrapper Feature Selection
Randall Wald, Taghi Khoshgoftaar and Amri Napolitano
Florida Atlantic University, USA

**Interpreting random forest models using a feature contribution method** ...................................................... 112
Anna Palczewska (1), Jan Palczewski (2), Richard Marchese Robinson (3) and Daniel Neagu (1)
(1) University of Bradford, UK
(2) University of Leeds, UK
(3) Syngenta Ltd., UK

**Entity Disambiguation in Tweets leveraging User Social Profiles** ................................................................. 120
Surender Reddy Yerva (1), Michele Catasta (1), Gianluca Demartini (2) and Karl Aberer (1)
(1) EPFL, Switzerland
(2) University of Fribourg, Switzerland

**Session A22: Databases, Structured/Semi-Structured Data I**

**Cross-Tenant Trust Models in Cloud Computing** .......................................................................................... 129
Bo Tang and Ravi Sandhu
University of Texas at San Antonio, USA

**PEP-Side Caching: An Insider Threat Port** ................................................................................................. 137
Qussai Yaseen (1), Qutaibah Althebyan (2) and Yaser Jararweh (2)
(1) Yarmouk University, Jordan
(2) Jordan University of Science and Technology, Jordan

**Web-based Metadata Retrieval Tool for Fine Art and Games Artwork** ......................................................... 145
Adam Aboutaleb (1), Riham Khalil (1), David Stockton (1) and Sheona Jackson (2)
(1) De Montfort University, UK
(2) Idhammar Systems Ltd., UK

**VFDS: Very Fast Database Sampling System** ............................................................................................... 153
Teodora Sandra Buda (1), Thomas Cerqueus (1), John Murphy (1) and Morten Kristiansen (2)
(1) University College Dublin, Ireland
(2) IBM Software Group, Dublin, Ireland

**Session A23: Special Session on Education and Special Session on Robotics and Intelligent Machines**

**A Vision of Education in 50 Years From Now** ............................................................................................... 161
June Rosemarie Massoud
School Boards in Montreal, Canada

**The possibilities of using Petri Nets for realization of a Universal Model of Educational Process** .......... 162
Zoltán Balogh, Milan Turčáni and Martin Mágdin
Constantine the Philosopher University in Nitra, Slovakia

**Geographic Clustering of Research Universities in Specialized Fields using Microsoft Academic Search** ......................................................................................................................... 170
David Kaczynski and Gongzhu Hu
Central Michigan University, USA

**Autonomy Requirements Engineering** ........................................................................................................ 175
Emil Vassev and Mike Hinchey
University of Limerick, Ireland
Service-Oriented Middleware for Collaborative UAVs ............................... 185
Nader Mohamed (1) and Jameela Al-Jaroodi (2)
(1) UAEU, Al Ain, UAE
(2) Middleware Technologies Lab., Bahrain

Session A31: Component-based Design and Reuse

Mining Reusable Software Components from Object-Oriented Source Code of a Set of Similar Software ................................................................. 193
Anas Shatnawi and Abdelhak Seriai
Université Montpellier 2 Sciences et Techniques, France

Detecting Distributed Software Components That Will Not Cause Emergent Behavior in Asynchronous Communication Style ........................................... 201
Fatemeh Hendijani Fard and Behrouz H. Far
University of Calgary, Canada

Feature-to-Code Traceability in a Collection of Software Variants: Combining Formal Concept Analysis and Information Retrieval ........................................ 209
Hamzeh Eyal-Salman, Abdelhak-Djamel Seriai and Christophe Dony
LIRMM Laboratory, France

A Component-Based Language Specific to Complex Systems Modeling .......... 217
Thouraya Bouabana-Tebibel (1), Stuart H. Rubin (2), Kadaouia Habib (1), Sofia Mellah (1) and Lynda Allata (1)
(1) Ecole Nationale Supérieure d’Informatique – Algiers, ALGERIA
(2) SPAWAR Systems Center Pacific – San Diego, USA

Session A32: Workshop on Information Reuse and Integration in Health Informatics (IRI-HI)

Comparison of Rank-Based vs. Score-Based Aggregation for Ensemble Gene Selection ............................................................ 225
David Dittman, Taghi Khoshgoftaar, Randall Wald and Amri Napolitano
Florida Atlantic University, USA

Hidden Dependencies between Class Imbalance and Difficulty of Learning for Bioinformatics Datasets ........................................................................... 232
Randall Wald, Taghi Khoshgoftaar, Alireza Fazelpour and David Dittman
Florida Atlantic University, USA

Radiation Therapy Simulation and Optimization Using Kinetic Polygon Modeling ................................................................. 239
David Allen and Ovidiu Daescu
The University of Texas at Dallas, USA

Automatic extraction of effective rule sets for Obstructive Sleep Apnea detection for a real-time mobile monitoring system ............................................ 247
Giovanna Sannino (1, 2), Ivanoe De Falco (1) and Giuseppe De Pietro (1)
(1) Institute of High Performance Computing and Networking, Italy
(2) University of Naples Parthenope, Italy

Patient Response Datasets: Challenges and Opportunities .......................... 254
Randall Wald and Taghi Khoshgoftaar
Florida Atlantic University, USA
Session B11: Artificial Intelligence and Optimization

Improving the Accomplishment of a Neural Network Based Agent for Draughts that Operates in a Distributed Learning Environment ................................................................. 262
Lidia B. P. Tomaz, Rita M. S. Julia and Ayres R. A. Barcelos
Federal University of Uberlândia, Brazil

Optimal Retrieval Algorithms for Personalized Content Aggregation ................................................. 270
Dan He and Douglass Parker
(i) IBM T.J. Watson Research, USA
(2) UCLA, USA

About logic-based A.I. systems that must handle more informative subsumed knowledge .............. 278
Eric Gregoire
Université d’Artois, France

Extracting and enriching workflows from text ..................................................................................... 285
Pol Schumacher, Mirjam Minor and Eric Schulte-Zurhausen
Goethe University Frankfurt, Germany

Session B12: Systems and Process Engineering

Siphon-Based Deadlock Prevention Policy for Interorganizational WorkFlow net Design ................... 293
Luciane Silva (1), Lígia Passos (2), Michel Soares (1) and Stéphane Julia (1)
(1) Federal University of Uberlândia, Brazil
(2) Federal Rural University of Rio de Janeiro, Brazil

CRIS - Computational Research Infrastructure for Science ................................................................. 301
Eduard Dragut, Peter Baker, Jia Xu, Muhammad Sarfraz, Elisa Bertino, Amgad Madhkour, Raghu Agarwal, Ahmed Mahmoud and Sangchun Han
Purdue University, USA

Integrating Information from Multiple Layers to Detect Data Exfiltration ........................................ 309
Puneet Sharma, Anupam Joshi and Tim Finin
University of Maryland, USA

Session B13: Biomedical & Healthcare Systems

Feature List Aggregation Approaches for Ensemble Gene Selection on Patient Response Datasets ... 317
Taghi Khoshgoftaar, Randall Wald, David Dittman and Amri Napolitano
Florida Atlantic University, USA

The Use of Balance-Aware Subsampling for Bioinformatics Datasets ............................................... 325
Randall Wald, Taghi Khoshgoftaar and Alireza Fazelpour
Florida Atlantic University, USA

Towards a Minimum Description Length Based Stopping Criterion for Semi-Supervised Time Series Classification .................................................................................................................... 333
Nurjahan Begum (1), Bing Hu (1), Thanawin Rakthanmanon (2) and Eamonn Keogh (1)
(1) University of California, Riverside, USA
(2) Kasetsart University, Thailand

Gene Selection Stability’s Dependence on Dataset Difficulty ............................................................... 341
David Dittman, Taghi Khoshgoftaar, Randall Wald and Amri Napolitano
Session B21: Cloud Computing and Large Scale Data and System Integration

An Integrated Cloud-Based Platform for Labor Monitoring and Data Analysis in Precision Agriculture
Li Tan, Ronald Haley, Riley Wortman, Yiannis Ampatzidis and Matthew Whiting
Washington State University, USA

An Automated Approach to Create, Store, and Analyze Large-scale Experimental Data in Clouds
Deepal Jayasinghe, Josh Kimball, Siddharth Choudhary, Tao Zhu and Calton Pu
Georgia Institute of Technology, USA

Cloud-Based Tasking, Collection, Processing, Exploitation, and Dissemination
Stuart H. Rubin (1) and Gordon K. Lee (2)
(1) SSC-PAC, San Diego, USA
(2) San Diego State University, USA

Efficient Live Wide Area VM Migration with IP Address Change Using Type II Hypervisor
Samuel A. Ajila and Omhennimhen Iyamu
Carleton University, Canada

Session B22: Automation and Distributed Systems

An Automated Ontology Generation Technique for an Emergent Behavior Detection System
Beenish Khurshid (1), Mohammad Moshirpour (1), Armin Eberlein (2) and Behrouz Far (1)
(1) University of Calgary, Canada
(2) American University of Sharjah, UAE

An Effective Method for Optimistic Message Logging Based On Group Communication Links
Jinho Ahn
Kyonggi University, South Korea

From Software Product Lines to System of Systems: Analysis of an evolution path
Marcelo Ramos (1), Rosana Braga (1), Paulo Masiero (1) and Rosângela Penteado (2)
(1) University of Sao Paulo, Brazil
(2) Federal University of Sao Carlos, Brazil

Efficiently Integrating MapReduce-based Computing into a Hurricane Loss Projection Model
Fausto Fleites (1), Steve Cocke (2), Shu-Ching Chen (1) and Shahid Hamid (1)
(1) Florida International University, USA
(2) Florida State University, USA

Session B23: Data Mining II

A New Energy Efficient Cluster-Head and Backup Selection Scheme in WSN
Davood Izadi, Jemal Abawajy and Sara Ghanavati
Deakin University, Australia

Filter- and Wrapper-Based Feature Selection for Predicting User Interaction with Twitter Bots
Randall Wald, Taghi Khoshgoftaar and Amri Napolitano
Florida Atlantic University, USA

A Survey of Stability Analysis of Feature Subset Selection Techniques
Probability-Based Incremental Association Rule Discovery Using the Normal Approximation

An Integrated Framework for Personalized Internet Workflows

Session C11: Reuse in Software Engineering I

A Model-based Repository of Security and Dependability Patterns for Trusted RCES

Porting Mobile Games in an Aspect-Oriented Way: An Industrial Case Study

The Dynamic Aspects of Product Derivation in DSPL: a Systematic Literature Review

Session C12: Machine Learning

A Reinforcement Learning-Based Algorithm for Deflection Routing in Optical Burst-Switched Networks

Behavioral Sequence Prediction for Evolving Data Stream

Mining Software Repositories to Acquire Software Risk Knowledge

Author Attribution on Streaming Data

Session C13: Workshop on Formal Methods Integration (FMI)

Learning-based Routing in MobileWireless Sensor Networks: Applying Formal Modeling and Analysis
**Fatemeh Kazemeyni** (1), **Olaf Owe** (1), **Einar Broch Johnsen** (1) and **Ilango Balasingham** (2)

(1) University of Oslo, Norway
(2) Oslo University Hospital, Norway

**Giving Alloy a family**

Renato Neves (1), Alexandre Madeira (2), Manuel A. Martins (2) and Luis Soares Barbosa (1)

(1) Univ. Minho, Portugal
(2) Univ. Aveiro, Portugal

**Creating GUI-based DSL formal tools**

Robson Silva (1), Alexandre Mota (1) and Rodrigo Rizzi Starr (2)

(1) Federal University of Pernambuco, Brazil
(2) São José dos Campos - SP, Brazil

**Problem Decomposition and Sub-Model Reconciliation of Control Systems in Event-B**

Sanaz Yeganefard and Michael Butler

University of Southampton, UK

**Verifying System C with Predicate Abstraction: a Component Based Approach**

Nesrine Harrath (1), Bruno Monsuez (1) and Kamel Barkaoui (2)

(1) ENSTA ParisTech, France
(2) CNAM Paris, France

**Session C21: Databases, Structured/Semi-Structured Data II**

**Parse Tree Based Approach for Processing XML Streams**

Mohammad Khabbaz (1), Dirar Assi (1), Reda Alhajj (1, 2) and Moustafa Hammad (3)

(1) University of Calgary, Canada
(2) Global University, Lebanon
(3) Google Inc., USA

**The FUSE Domain-Aware Approach to User Model Interoperability: A Comparative Study**

Eddie Walsh, Alexander O'Connor and Vincent Wade

Trinity College Dublin, Ireland

**A Fast and Compact Indexing Technique for Moving Objects**

Yonghun Park (1), Ling Liu (1) and Jaesoo Yoo (2)

(1) Georgia Institute of Technology, USA
(2) Chungbuk National University, Korea

**Information Reuse in Dynamic Spectrum Access**

Prashant Krishnamurthy, Martin Weiss and David Tipper

University of Pittsburgh, USA

**Session C22: Reuse in Software Engineering II**

**A Brechó-EcoSys Extension to Support Negotiation in the Software Ecosystems Context**

Rodrigo Santos, Lucas Tostes and Claudia Werner

Federal University of Rio de Janeiro, Brazil

**Mining Features from the Object-Oriented Source Code of Software Variants by Combining Lexical and Structural Similarity**

Ra'Fat Al-Msie'Deen (1), Abdelhak Seriai (1), Marianne Huchard (1), Christelle Urtado (2) and Sylvain Vauttier (2)
An Approach to Develop Frameworks from Feature Models

Matheus Viana (1), Rosângela Penteado (1), Antônio Francisco Do Prado (1) and Rafael Durelli (2)

(1) Federal University of S~ao Carlos, Brazil
(2) University of S~ao Paulo, Brazil

Session C23: Multimedia Systems

Comparison of the structure and the computational complexity of the Inverse Difference and Laplacian pyramids for still image decomposition

Roumen Kountchev (1), Stuart Rubin (2), Mariofanna Milanova (3) and Roumiana Kountcheva (4)

(1) Technical University of Sofia, Bulgaria
(2) Space and Naval Warfare Systems Center, San Diego, USA
(3) UALR, USA
(4) Sofia, Bulgaria

Road Segment Partitioning towards Anomalous Trajectory Detection for Surveillance Applications

Muhammad Aamir Saleem, Waqas Nawaz, Youngkoo Lee and Sungyoung Lee

Kyung Hee University, South Korea

Adaptive Association Rule Mining for Web Video Event Classification

Chengde Zhang (1), Xiao Wu (1), Mei-Ling Shyu (2) and Qiang Peng (1)

(1) Southwest Jiaotong University, China
(2) University of Miami, USA

Spatial-temporal Motion Information Integration for Action Detection and Recognition in Non-static Background

Dianting Liu (1), Mei-Ling Shyu (1) and Guiru Zhao (2)

(1) University of Miami, USA
(2) China Earthquake Networks Center, China

Session C31: Government/Industry/Application Track and Special Session on Power and Energy

A Knowledge-Based Semantic Tool for Standard Part Management in Aerospace Industry

Hasan Mert Taymaz (1), Tansel Özyer (2) and Can Cangelir (1)

(1) Turkish Aerospace Industry TAI, Turkey
(2) University Ankara, Turkey

Simulation study on evacuation from buildings

Robert Kosiński (1, 2) and Andrzej Grabowski (2)

(1) Warsaw University of Technology, Poland
(2) Central Institute for Labour Protection – National Research Institute, Poland

Low Power-Energy Storage System for Energy Harvesting Applications

Russ Tatro, Cody Tudor, Eric Sprung and Justin Meyer

California State University, Sacramento, USA

Low Power Design of a Wireless Sensor Node for Indoor Monitoring
Russ Tatro, Eric Sprung, Cody Tudor and Justin Meyer  
*California State University, Sacramento, USA*

**Session C32: Workshop on Data Integration and Mining (DIM)**

**Cultural Considerations for Designing Crossover Applications for the Visually Impaired**

Abdullah Almurayh and Sudhanshu Semwal  
*University of Colorado Colorado Springs, USA*

**Controlling Xen Cloud Platform via Smart Phones**

Abdullah Almurayh and Sudhanshu Semwal  
*University of Colorado Colorado Springs, USA*

**A Comparative Study of Data Mining Techniques for Credit Scoring in Banking**

Shih-Chen Huang and Min-Yuh Day  
*Tamkang University, Taiwan*

**MFWK-Means: Minkowski Metric Fuzzy Weighted K-Means for high dimensional data clustering**

Liudmila Svetlova (1, 2), Boris Mirkin (3) and Hansheng Lei (2)  
(1) *Moscow Institute of Physics and Technology, RF*  
(2) *University of Texas at Brownsville, USA*  
(3) *University of London, UK*

**Big Data Analysis Using Apache Hadoop**

Jyoti Nandimath, Saumitra Vaidya, Ankur Patil, Ekata Banerjee and Pratima Kakade  
*SKNCOE, Pune, India*

**Session SP: Poster Exhibition**

**A linear filtering theory-based approach for load shedding**

Lorena Chavarría-Báez and Rosaura Palma-Orozco  
*Istituto Politécnico Nacional, Mexico*

**Reinforcement of Trust in DSR**

Fatma Laidoui (1) and Thouraya Bouabana-Tebibel (2)  
(1) *École nationale Supérieure d’Informatique – ESI, Algeria*  
(2) *Ecole Militaire Polytechnique – EMP, Algeria*

**Federated Ensemble Kalman Filter in No Reset Mode Design**

Maryam Kazerooni (1), Faridoon Shabaninia (1), Mohammad Vaziri (2) and Suresh Vadhva (2)  
(1) *Shiraz University, Iran*  
(2) *California State University, Sacramento, USA*

**ReadFast: Optimizing Structural Search Relevance for Big Biomedical Text**

Michael Gubanov and Anna Pyayt  
(1) *Massachusetts Institute of Technology, USA*  
(2) *University of South Florida, USA*

**Using online stream-processing for portable electroencephalography system for stress and fear detect**

Surya Cheemalapati (1), Michael Gubanov (2), Michael Del Vale (1) and Anna Pyayt (1)  
(1) *University of South Florida, USA*  
(2) *Massachusetts Institute of Technology, USA*
Natural Image Segmentation Using Morphological Mathematics and Fuzzy Logic ............................................. 724
Victoria Lynn Fox and Mariofana Milanova
University of Arkansas at Little Rock, USA

Understanding Environmental Influences on Performing Password-based Mobile Authentication ........ 728
Sergey V. Maydebura, Dong Hyun Jeong and Byunggu Yu
University of the District of Columbia, USA

Distributed Ensemble Kalman Filter for Multisensor Application ................................................................. 732
Maryam Kazerooni (1), Faridoon Shabaninia (1), Mohamad Vaziri (2) and Suresh Vadhva (2)
(1) Shiraz University, Iran
(2) California State University, Sacramento, USA

Author Index ....................................................................................................................................................... 737