



IMMM 2015

The Fifth International Conference on Advances in Information Mining and
Management

DATASETS 2015

The International Symposium on Challenges for Designing and Using Datasets

June 21 - 26, 2015

Brussels, Belgium

IMMM 2015 Editors

Christian Ernst, Ecole des Mines de St Etienne - Gardanne, France
Andreas Schmidt, Karlsruhe University of Applied Sciences, Germany

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2015) by International Academy, Research, and Industry Association (IARIA)
Please refer to the Copyright Information page.

Printed by Curran Associates, Inc. (2015)

International Academy, Research, and Industry Association (IARIA)
412 Derby Way
Wilmington, DE 19810

Phone: (408) 893-6407
Fax: (408) 527-6351

petre@iaria.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

TABLE OF CONTENTS

IMMM 1: INFORMATION MINING AND MANAGEMENT

Closed Frequent Itemset Mining over Fast Data Stream Based on Hadoop	1
<i>Shan Jicheng, Liu Qingbao</i>	
Robustness of Bisecting k-means Clustering-based Collaborative Filtering Algorithm	7
<i>Alper Bilge, Huseyin Polat</i>	
Improving Relevance Effectiveness in Data Leakage Detection Using Feature Selection	14
<i>Adrienn Skrop</i>	
Analyzing and Improving Educational Process Models using Process Mining Techniques	17
<i>Awatef Hicheur Cairns, Billel Gueni, Joseph Assu, Christian Joubert, Nasser Khelifa</i>	

IMMM 2: AUTOMATED RETRIEVAL AND MINING

Streamlining the Detection of Accounting Fraud through Web Mining and Interpretable Internal Representations	23
<i>Duarte Trigueiros, Carolina Sam</i>	
Exponential Moving Maximum Filter for Predictive Analytics in Network Reporting	27
<i>Bin Yu, Les Smith, Mark E. Threefoot</i>	
Automatic KDD Data Preparation Using Multi-criteria Features	33
<i>Youssef Hmamouche, Christian Ersnt, Alain Casali</i>	
Towards Predictive Policing: Knowledge-based Monitoring of Social Networks	39
<i>Michael Spranger, Florian Heinke, Steffen Grunert, Dirk Labudde</i>	

IMMM 3: DATA MANAGEMENT IN SPECIAL ENVIRONMENTS

Real-time Partition of Streamed Graphs for Data Mining Over Large Scale Data	41
<i>Victor Medel, Unai Arronategui</i>	
Sketch of Big Data Real-Time Analytics Model	48
<i>Bakhtiar Amen, Joan Lu</i>	
Augmenting Data Files with Semantics for Coherency, Extensibility, and Reproducibility	54
<i>John McCloud, Subhasish Mazumdar</i>	
Automatically Triggering Activity and Product Predictions in Mobile Phone Based on Individual's Activity	61
<i>Kalpana Algotar, Sanjay Addicam</i>	

DATASETS 1: DATASET USE AND DESIGN

Recommender Systems for Museums: Evaluation on a Real Dataset	65
<i>Ivan Keller, Emmanuel Viennet</i>	
An Extensible Conceptual Model for Tabular Scientific Datasets	72
<i>Javad Chamanara, Michael Owonibi, Alsayed Algergawy, Roman Gerlach</i>	
Context-aware Healthcare Dataset- A Case Study from Pakistan	77
<i>Shahid Mahmud, Rahat Iqbal, Faiyaz Doctor</i>	
Query Acceleration in Multimedia Database Systems	83
<i>Ramzi Haraty, Rawa Karaki</i>	
Author Index	