

EDM 2012

**5th International Conference on
Educational Data Mining**

PROCEEDINGS OF THE
FIFTH INTERNATIONAL CONFERENCE ON
EDUCATIONAL DATA MINING

Chania, Greece, June 19-21 2012

**Kalina Yacef, Osmar Zaïane,
Arnon HersHKovitz, Michael Yudelson, John Stamper**

Kalina Yacef, Osmar Zaïane, Arnon HersHKovitz, Michael Yudelson and John Stamper

International Conference on Educational Data Mining (EDM) 2012
Proceedings of the 5th International Conference on Educational Data Mining
Kalina Yacef, Osmar Zaïane, Arnon HersHKovitz, Michael Yudelson and John Stamper (eds.)
Chania, June 19-21, 2012

ISBN: 978-1-74210-276-4

Preface

The 5th International Conference on Educational Data Mining (EDM 2012) is held in picturesque Chania on the beautiful Crete island in Greece, under the auspices of the International Educational Data Mining Society (IEDMS). The EDM 2012 conference is a leading international forum for high quality research that mines large data sets of educational data to answer educational research questions. These data sets may come from learning management systems, interactive learning environments, intelligent tutoring systems, or any system used in a learning context.

We received a total of 50 full papers, 19 short papers and 20 posters from authors of 30 different countries and these submissions covered the full range of the conference topics specified in the call for papers. The papers were distributed to the 56 international program committee members, who called upon additional 17 external reviewers. All papers were reviewed by 3 reviewers and any paper submitted by organisers was sent to anonymous reviewers unbeknownst to the involved organisers. After receiving a total of 235 reviews, 17 full papers and 15 short papers were accepted to be presented at the conference and included in these proceedings, giving an acceptance rate of 34% for full papers, and 46% overall. Furthermore, 17 posters were accepted and their summaries appear in the proceedings. We have also included the abstracts of the 3 invited talks by (i) Professor Myra Spiliopoulou, Professor of Business Information Systems, Computer Science, Otto-von-Guericke University Magdeburg, Germany, (ii) Professor Danielle S. McNamara, Learning Sciences Institute Psychology Department, Arizona State University and (iii) Dr Bob Dolan, Senior Research Scientist, Assessment and Information, Pearson.

We would like to thank our sponsors Carnegie Learning (Gold Level), Pearson (Gold Level) and LearnLab (Silver Level) for their generous support. We would also like to thank the program committee members, the additional reviewers, the local committee, the web chair and the invited speakers for their invaluable help in putting this program and conference together.

We hope that the papers contained in these proceedings will be stimulating to most of the readers, provide thought-provoking new ideas to motivate new research, and prove helpful towards improving the literacy in educational data mining.

Enjoy the conference.

Kalina Yacef
Osmar Zaiane
Arnon HersHKovitz
Michael Yudelson
John Stamper

Organization

CONFERENCE CHAIR

John Stamper Carnegie Mellon University, USA

PROGRAM CHAIRS

Kalina Yacef The University of Sydney, Australia
Osmar Zaiane University of Alberta, Canada

POSTER AND DEMO CHAIRS

Arnon HersHKovitz Worcester Polytechnic Institute, USA
Michael Yudelson Carnegie Mellon University, USA

FINANCE AND WEB CHAIR

Michael Bett Carnegie Mellon University, USA

LOCAL ORGANIZING TEAM

Kitty Panourgia Neoanalysis, Greece

STEERING COMMITTEE

Ryan Baker Worcester Polytechnic Institute, USA
Tiffany Barnes University of North Carolina at Charlotte, USA
Joseph E. Beck Worcester Polytechnic Institute, USA
Michel Desmarais Ecole Polytechnique de Montreal, Canada
Neil Heffernan Worcester Polytechnic Institute, USA
Agathe Merceron Beuth University of Applied Sciences, Germany
Mykola Pechenizkiy Eindhoven University of Technology, Netherlands
Cristobal Romero Cordoba University, Spain
Kalina Yacef The University of Sydney, Australia

PROGRAM COMMITTEE

Eσμα Aimeur University of Montreal, Canada
Mirjam Augstein Kock Upper Austria University of Applied Sciences, Austria
Elizabeth Ayers University of California, USA
Tiffany Barnes University of North Carolina at Charlotte, USA
Joseph Beck Worcester Polytechnic Institute, USA
Bettina Berendt Katholieke Universiteit Leuven, Belgium
Gautam Biswas Vanderbilt University, USA
Jesus G. Boticario U.N.E.D., Spain
Rafael A. Calvo The University of Sydney, Australia
Min Chi Stanford University, USA
Christophe Choquet University of Maine, USA
Richard Cox University of Edinburgh, Scotland
Sidney D'Mello The University of Memphis, USA
Michel Desmarais Ecole Polytechnique de Montreal, Canada
Mingyu Feng SRI International, USA

Davide Fossati	Georgia Institute of Technology, USA
Eva Gibaja	Cordoba University, Spain
Daniele Godoy	Universidad Nacional del Centro de la Pcia, Buenos Aires, Argentina
Neil Heffernan	Worcester Polytechnic Institute, USA
Arnon HersHKovitz	Worcester Polytechnic Institute, USA
Roland Hubscher	Bentley University, USA
Sebastien Iksal	LIUM - University of Maine Computer Science Laboratory, Le Mans, France
Judy Kay	The University of Sydney, Australia
Jihie Kim	University of Southern California, USA
Evgeny Knutov	Technische Universiteit Eindhoven, The Netherlands
Kenneth Koedinger	Carnegie Mellon University, USA
Irena Koprinska	The University of Sydney, Australia
Vanda Luengo	Université Joseph Fourier, Grenoble, France
Tara Madhyastha	University of Washington, USA
Noboru Matsuda	Carnegie Mellon University, USA
Manolis Mavrikis	University of London, UK
Riccardo Mazza	University of Lugano/University of Applied Sciences of Southern Switzerland, Switzerland
Gordon Mccalla	University of Saskatchewan, Canada
Bruce McLaren	Carnegie Mellon University, USA
Agathe Merceron	Beuth University of Applied Sciences Berlin, Germany
Julià Minguillón	Universitat Oberta de Catalunya, Spain
Jack Mostow	Carnegie Mellon University, USA
Kasia Muldner	Arizona State University, USA
Roger Nkambou	Université du Québec À Montréal, Canada
Alvaro Ortigosa	Universidad Autonoma de Madrid, Spain
Alexandros Paramythis	Johannes Kepler University, Austria
Zachary Pardos	Worcester Polytechnic Institute, USA
Mykola Pechenizkiy	Eindhoven University of Technology, The Netherlands
Cristobal Romero	Cordoba University, Spain Carolyn
Rose	Carnegie Mellon University, USA
Ryan S.J.D. Baker	Worcester Polytechnic Institute, USA
Erin Shaw	University of Southern California, USA
John Stamper	Carnegie Mellon University, USA
Jun-Ming Su	National Chiao Tung University, Taiwan
Steven Tanimoto	University of Washington, USA
Sebastian Ventura	Cordoba University, Spain
Stephan Weibelzahl	National College of Ireland, Ireland
Kalina Yacef	The University of Sydney, Australia
Michael Yudelson	Carnegie Mellon University, USA
Amelia Zafra Gómez	Cordoba University, Spain
Osmar R. Zaiane	University of Alberta, Canada

ADDITIONAL REVIEWERS

Iiya Goldin

I-Han Hsiao

Nan Li

Ming Liu

Collin Lynch

Roberto Martinez Maldonado

Terry Peckham

Hamman Samuel

Michael Sao Pedro

Mayya Sharipova

Carlos Soares

Mansoureh Takaffoli

Eric Verbeek

Qingyang Wang

IEDMS OFFICERS AND BOARD OF DIRECTORS

Ryan Baker (President)

Worcester Polytechnic Institute, USA

Michel Desmarais (Treasurer)

Ecole Polytechnique de Montreal, Canada

Tiffany Barnes

University of North Carolina at Charlotte, USA

Joseph E. Beck

Worcester Polytechnic Institute, USA

Neil Heffernan

Worcester Polytechnic Institute, USA

Agathe Merceron

Beuth University of Applied Sciences, Germany

Mykola Pechenizkiy

Eindhoven University of Technology, Netherlands

Cristobal Romero

Cordoba University, Spain

Kalina Yacef

The University of Sydney, Australia

Sponsors



PEARSON



<http://www.carnegielearning.com>

<http://www.pearson.com>

<http://www.learnlab.org>

<http://www.educationaldatamining.org>

Table of contents

Invited Talks (abstracts)

Stream Mining in Education? Dealing with Evolution	3
<i>Professor Myra Spiliopoulou</i>	
From Text to Feedback: Leveraging Data Mining to Build Educational Technologies	5
<i>Danielle S. McNamara</i>	
Five Aspirations for Educational Data Mining	7
<i>Bob Dolan and John Behrens</i>	

Full Papers

Assisting Instructional Assessment of Undergraduate Collaborative Wiki and SVN Activities	10
<i>Jihie Kim, Erin Shaw, Hao Xu and Adarsh G V</i>	
Automated Student Model Improvement	17
<i>Kenneth R. Koedinger, Elizabeth A. McLaughlin and John C. Stamper</i>	
Automatic Discovery of Speech Act Categories in Educational Games	25
<i>Vasile Rus, Arthur Graesser, Cristian Moldovan and Nobal Niraula</i>	
Co-Clustering by Bipartite Spectral Graph Partitioning for Out-of-Tutor Prediction	33
<i>Shubhendu Trivedi, Zachary Pardos, Gábor Sárközy and Neil Heffernan</i>	
Comparison of methods to trace multiple subskills: Is LR-DBN best?	41
<i>Yanbo Xu and Jack Mostow</i>	
Dynamic Cognitive Tracing: Towards Unified Discovery of Student and Cognitive Models	49
<i>José González-Brenes and Jack Mostow</i>	
Identifying Learning Behaviors by Contextualizing Differential Sequence Mining with Action Features and Performance Evolution	57
<i>John Kinnebrew and Gautam Biswas</i>	
Identifying Students' Characteristic Learning Behaviors in an Intelligent Tutoring System Fostering Self-Regulated Learning	65
<i>François Bouchet, John Kinnebrew, Gautam Biswas and Roger Azevedo</i>	

Learner Differences in Hint Processing	73
<i>Ilya Goldin, Kenneth Koedinger and Vincent Aleven</i>	
Methods to find the number of latent skills	81
<i>Behzad Beheshti, Michel Desmarais and Rhouma Naceur</i>	
Mining Student Behavior Patterns in Reading Comprehension Tasks	87
<i>Terry Peckham and Gordon McCalla</i>	
Model-Based Collaborative Filtering Analysis of Student Response Data: Machine-Learning Item Response Theory	95
<i>Yoav Bergner, Stefan Droschler, Gerd Kortemeyer, Saif Rayyan, Daniel Seaton and David Pritchard</i>	
Predicting drop-out from social behaviour of students	103
<i>Tomas Obsivac, Lubos Popelinsky, Jaroslav Bayer, Jan Geryk and Hana Bydzovska</i>	
Searching for Variables and Models to Investigate Mediators of Learning from Multiple Representations	110
<i>Martina Rau and Richard Scheines</i>	
The Impact on Individualizing Student Models on Necessary Practice Opportunities	118
<i>Jung In Lee and Emma Brunskill</i>	
Towards Sensor-Free Affect Detection in Cognitive Tutor Algebra	126
<i>Ryan S.J.D. Baker, Sujith Gowda, Michael Wixon, Jessica Kalka, Angela Wagner, Aatish Salvi, Vincent Aleven, Gail Kusbit, Jaclyn Ocumpaugh and Lisa Rossi</i>	
Using Edit Distance to Analyse Errors in a Natural Language to Logic Translation Corpus	134
<i>Dave Barker-Plummer, Robert Dale, Richard Cox and Alex Romanczuk</i>	
 Short Papers	
Calculating Probabilistic Distance to Solution in a Complex Problem Solving Domain	144
<i>Leigh Ann Sudol, Kelly Rivers and Thomas K. Harris</i>	
Classification via clustering for predicting final marks starting from the student participation in Forums	148
<i>Manuel Ignacio Lopez, Cristobal Romero, Sebastián Ventura and J.M. Luna</i>	
Development of a Workbench to Address the Educational Data Mining Bottleneck	152
<i>Ma. Mercedes Rodrigo, Ryan S. J. D. Baker, Bruce McLaren, Alejandra Jayme and Thomas Dy</i>	

Early Prediction of Student Self-Regulation Strategies by Combining Multiple Models	156
<i>Jennifer Sabourin, Bradford Mott and James Lester</i>	
Identifying Successful Learners from Interaction Behaviour	160
<i>Judi Mccuaig and Julia Baldwin</i>	
Interaction Networks: Generating High Level Hints Based on Network Community Clusterings	164
<i>Michael Eagle, Matthew Johnson and Tiffany Barnes</i>	
Interleaved Practice with Multiple Representations: Analyses with Knowledge Tracing Based Techniques	168
<i>Martina Rau and Zachary Pardos</i>	
Learning Gains for Core Concepts in a Serious Game on Scientific Reasoning	172
<i>Carol Forsyth, Philip Pavlik Jr, Arthur Graesser, Zhiqiang Cai, Mae-Lynn Germany, Keith Millis, Heather Butler, Diane Halpern and Robert Dolan</i>	
Leveraging First Response Time into the Knowledge Tracing Model	176
<i>Yutao Wang and Neil Heffernan</i>	
Meta-learning Approach for Automatic Parameter Tuning: A case of study with educational datasets	180
<i>María De Mar Molina, Cristobal Romero, Sebastián Ventura and J.M. Luna</i>	
Mining Concept Maps to Understand University Students' Learning	184
<i>Jin Soung Yoo and Moon-Heum Cho</i>	
Policy Building – An Extension To User Modeling	188
<i>Michael Yudelson and Emma Brunskill</i>	
The real world significance of performance prediction	192
<i>Zachary Pardos, Qing Yang Wang and Shubhendu Trivedi</i>	
The Rise of the Super Experiment	196
<i>John Stamper, Derek Lomas, Dixie Ching, Steven Ritter, Kenneth Koedinger and Jonathan Steinhart</i>	
Using Student Modeling to Estimate Student Knowledge Retention	200
<i>Yutao Wang and Joseph Beck</i>	
Posters	
A promising classification method for predicting distance students' performance.	206
<i>Diego García-Saiz and Marta Zorrilla</i>	

Analyzing paths in a student database	208
<i>Donatella Merlini, Renza Campagni and Renzo Sprugnoli</i>	
Analyzing the behavior of a teacher network in a Web 2.0 environment	210
<i>Eliana Scheihing, Carolina Aros and Daniel Guerra</i>	
Automated Detection of Mentors and Players in an Educational Game	212
<i>Fazel Keshtkar, Brent Morgan and Arthur Graesser</i>	
Categorizing Students' Response Patterns using the Concept of Fractal Dimension	214
<i>Rasil Warnakulasooriya and William Galen</i>	
CurriM: Curriculum Mining	216
<i>Mykola Pechenizkiy, Nikola Trcka, Paul De Bra and Pedro Toledo</i>	
Data mining techniques for design of ITS student models	218
<i>Ritu Chaturvedi and Christie Ezeife</i>	
Deciding on Feedback Polarity and Timing	220
<i>Stuart Johnson and Osmar Zaiane</i>	
Finding Dependent Test Items: An Information Theory Based Approach	222
<i>Xiaoxun Sun</i>	
Fit-to-Model Statistics for Evaluating Quality of Bayesian Student Ability Estimation	224
<i>Ling Tan</i>	
Inferring learners' knowledge from observed actions	226
<i>Anna Rafferty, Michelle Lamar and Thomas Griffiths</i>	
Learning Paths in a Non-Personalizing e-Learning Environment	228
<i>Agathe Merceron, Liane Beuster, Margarita Elkina, Albrecht Fortenbacher, Leonard Kappe, Andreas Pursian, Sebastian Schwarzrock and Boris Wenzlaff</i>	
Similarity Functions for Collaborative Master Recommendations	230
<i>Alexandru Surpatean, Evgueni Smirnov and Nicolai Manie</i>	
Social Networks Analysis for Quantifying Students' Performance in Teamwork	232
<i>Pedro Crespo and Claudia Antunes</i>	
Speaking (and touching) to learn: a method for mining the digital footprints of face-to-face collaboration	234
<i>Roberto Martinez Maldonado, Kalina Yacef and Judy Kay</i>	
Stress Analytics in Education	236
<i>Rafal Kocielnik, Mykola Pechenizkiy and Natalia Sidorova</i>	

Variable Construction and Causal Discovery for Cognitive Tutor Log Data: Initial Results	238
<i>Stephen Fanesali</i>	

List of Authors	241
------------------------	------------

