
Introduction

Reda Alhajj*

Canada Department of Computer Science,
University of Calgary,
Alberta, Canada
E-mail: alhajj@ucalgary.ca
*Corresponding author

Kang Zhang

Department of Computer Science,
Erik Jonsson School of Engineering and Computer Science,
The University of Texas at Dallas,
P.O. Box 830688,
Mail Stop EC31,
Richardson, TX 75080 3021, USA
Fax: +1 972 883 2349
E-mail: kzhang@utdallas.edu

Biographical notes: Reda Alhajj is a Professor of Computer Science at the University of Calgary. He published over 280 papers in prestigious journals and conferences. He served on the program committee of several international conferences including IEEE ICDE, IEEE ICDM, IEEE IAT, SIAM DM, etc. He also served as a Guest Editor of several special issues and is currently the Program Chair of IEEE IRI 2009, CaSoN 2009, ASONAM 2009, ISCIS 2009, MS 2009 and OSINT-WM 2009. He is on the Editorial Board of several journals and Associate Editor of IEEE SMC – Part C. His research areas cover biocomputing and biodata analysis, data mining, multiagent systems, schema integration and re-engineering, social networks and XML. He received Outstanding Achievements in Supervision Award from the Faculty of Graduate Studies at the University of Calgary.

Kang Zhang is a Professor of Computer Science Department at the University of Texas at Dallas. His current research interests are in the areas of information visualisation, visual programming and visual languages, and web engineering; he has published over 180 papers in these areas. He has authored and edited four books. His research has been funded by the UK SERC, Australian Research Council, Sun Microsystems, Texas State, US NSF and US Department of Education. He has been the General Chair and Program Chair of several major international conferences. He is also on the Editorial Boards of *Journal of Visual Languages and Computing*, and *Int. J. Software Engineering and Knowledge Engineering*.

The increasing volumes and dimensions of information have dramatic impact on effective decision-making. To remedy this situation, *information reuse and integration* (IRI) seeks to maximise the reuse of information by creating simple, rich and reusable knowledge representations and consequently explores strategies for integrating this knowledge into

legacy systems. IRI plays a pivotal role in the capture, representation, maintenance, integration, validation and extrapolation of information; and applies both information and knowledge for enhancing decision-making in various application domains. This special issue includes five papers, mainly selected from the papers that were accepted and presented at the *2008 IEEE International Conference on Information Reuse and Integration* (IEEE IRI'2008), which was held in Las Vegas in July 2008. The conference attracted a large volume of high quality papers. Due to excellent reviews received at IRI'08, the authors of these papers were invited to submit significantly extended and enhanced versions to this special issue. After another round of reviews by the special issue review board, the five papers have been further revised and finally accepted to be included in this special issue.

The five papers cover an interesting range of topics including randomisation and its application to logic-based workflow, multidimensional index structures, data warehouse modelling and design, adaptive multiple super-page queue management, web services, accurate classifier ensemble and web-based collaborative information system.

The first paper by Kim describes a modelling and analysis of a web-based collaborative information system – Petri net-based collaborative enterprise. In his paper, Kim discusses four essential phases of designing a human-centric collaborative enterprise: framework, a case study, modelling and analysis. He first proposes an application-oriented human-centric collaborative commerce architecture; second, he developed an appropriate case study; third, he shows the modelling and simulation of the selected application; and finally he discusses some analysis of the application using Stochastic Petri nets.

The second paper by Liang and Rubin presents a rule randomisation and its application to logic-based workflow. The authors study how randomisation enables reuse and reduces processing time in logic-based workflow verification approaches. In particular, they look at a propositional logic-based workflow verification technique. For the logic inference rules, which are used to infer new truthful propositions from existing truthful propositions in this logic, they apply randomisation to the inference rules after each verification task such that new inference rules, reflecting the componentised verification, are added to the inference rule sets. They also reviewed the savings gained by verifying a workflow pattern and provide a theoretical analysis.

The third paper by Rifaie, Kianmehr, Alhadj and Ridley discusses data modelling for effective data warehouse architecture and design. The authors present the process of data warehouse architecture development and design. They highlight the different aspects to be considered in building a data warehouse. These range from data store characteristics to data modelling and the principles to be considered for effective data warehouse architecture.

The fourth paper by Su and Khoshgoftaar describes an approach for making an accurate classifier ensemble by voting on classifications from imputed learning sets. The authors present VCI as a system that produces those diverse training sets by randomly removing a small percentage of attribute values from the original training set, and then using an imputation technique to replace those values. VCI then runs a learning algorithm on each of these imputed training sets to produce a set of base classifiers. Later, the final prediction on a novel instance is the plurality classification produced by these classifiers. They investigate various imputation techniques. The reported results are very promising.

The fifth paper of Itshak and Wiseman introduces AMSQM as an adaptive multiple super-page queue management. The authors based their work on the fact that there are

some particular operating systems that support super-paging and there are some recent research papers that show interesting ideas on how to intelligently integrate them; however, nowadays Operating System's page replacement mechanism still uses the old Clock algorithm which gives the same priority to small and large pages. Based on this, the authors show a technique that enhances the page replacement mechanism to an algorithm based on more parameters and is suitable for a super-paging environment.

Finally, we would like to thank all the authors for their efforts in enhancing and extending their papers for this special issue. The Editor-In-Chief, Prof. John Wang's guidance and support are particularly appreciated. Thanks also goes to external reviewers for their valuable assistance in making this special issue published in a timely fashion.