Preface

Subhash Bhalla

Graduate School of Computer Science and Engineering,
University of Aizu,
Aizu-wakamatsu, Fukushima, 965-8580, Japan
Email: bhalla.subhash@gmail.com

Biographical notes: Subhash Bhalla is a Professor at the Graduate School of Computer Science and Engineering, University of Aizu. He obtained his PhD in Computer Science in 1984. His research interests include distributed information systems, real-time operating systems, new query languages for web data resources and information retrieval in healthcare.

Large-scale information systems in transportation, banking and public-utility services depend on computing infrastructure. Many research efforts are being made in related areas, such as web-based computing, information accesses by web users, wireless computing and sensor networks. Government agencies in many countries plan to launch facilities in education, healthcare and information support as a part of e-government initiatives. In this context, information interchange management has become an active research field. A number of new opportunities have evolved in design and modelling based on new computing needs of the users. Database systems play a central role in supporting networked information systems for access and storage management aspects.

In the same context, the 10th International Workshop on Databases in Networked Information Systems (DNIS) 2015 was held on March 23–25, 2015 at the University of Aizu in Japan. The workshop program included research contributions and invited contributions. These were focused on data analytics in sciences. Some of the manuscripts considered topics related to data models and new query languages, in big data analytics. A view of research activity in related areas was provided by a special session on the topics. These included big data analysis, information knowledge and management, business analytics and visualisation, networked information systems, and business data analytics in astronomy and sciences. A few of the invited contributions have been contributed by Florin Rusu, Punam Bedi, Shelly Sachdeva, Yasuhiko Morimoto and Lukas Pichl. Revised versions of these and other selected contributed manuscripts have been included in this special issue.

The articles in this issue begin with a paper by Zhao et al., introducing automatic identification and classification of astronomical objects in a proposed GLADE system. Recent developments in time-domain astronomy are connected with the Palomar Transient Factory project. It has a comprehensive detection system for the identification and classification of transient astrophysical objects. Its classification system consists of two components, namely real-time and offline. The entire volume of stream data is also archived for future references. The result of the proposed system is an efficient classifier implementation capable of processing a new set of acquired images in a short time. Second, the manuscript introduces a parallel similarity join algorithm for advanced transient classification.

In the information knowledge and management section, the paper by Bedi et al. considers a multi-agent recommender system with a focus on trust and reputation. It studies a complex case of user profile modelling. The study proposes a trust and reputation-based collaborative filtering algorithm. A multi-agent recommender system for e-tourism for recommending tourism services using the above algorithm is proposed. It can generate recommendations for hotels, places to visit and restaurants.

In the business analytics and visualisation section, the article by Batra and Sachdeva explores the multi-table entity attribute value data model. It considers storage level implementation for standardised electronic health records databases. This paper proposes a search-efficient avenue for storing a database containing specialised data. The proposal outperforms other existing models (entity attribute value model, dynamic tables, optimised entity attribute value and optimised column oriented model) under various query scenarios and varying dataset sizes. The paper also examines the maximum percentage of non-null density appropriate for choosing the proposed method as a storage option.

In the same business analytics and visualisation section, the paper by Zaman et al. considers k-objects selection for a keyword query. Keyword query interfaces are commonly used as a standard in information retrieval. The user gives a keyword, and objects that are closely related to that keyword are returned to the user. Top-k query is one of the popular methods to select important objects from a large number of candidates. During the selection of k-objects, the proposed method prevents disclosures of sensitive values. The manuscript outlines a process that uses parallelism using map-reduce.

In the subsequent section on networked information systems, the manuscript on vehicular ad hoc networks considers adaptive traffic control for efficient response. It uses a parallel computing platform. One of the goals of the

adaptive traffic control algorithm is to take care of conflicting movements at the intersections to avoid accidents and ensure smooth flow of traffic in a safe and efficient manner. Jindal and Bedi consider various load conditions, by varying the numbers of vehicles for considering real world networks. The performance of the algorithm has been compared with a serial counterpart.

The following article in this section, by Chu, considers workflow patterns for qualitatively aggregation of information. The system considers a network with distributed public information resources. The proposed system aggregates information from utility companies, bus companies, city information from the internet, and the public facilities map of the city, to generate geographic data. Many different query methods are used to obtain the target information. The query results and the shortest bus route are visualised on the embedded map. The detailed search information is shown in the side-bar. The system is helpful for city residents and city visitors.

In the section on business data analytics in astronomy and sciences, the manuscript by Pichl and Kaizoji considers data mining of causality extent in time-series, in the domain of economic data. It is an important area of computational intelligence research. It is used in algorithmic trading or risk diversification strategies. The authors adopt the support vector machine (SVM) and artificial neural network (ANN) for causality rate extraction. They consider the data of the commodity of gasoline traded in Japan. The trend prediction analysis is complemented by the ANN method with four hidden layers.

In the concluding section, the paper by Ceriani and Bottoni considers linked data with a dataflow platform and applications. Its modelling considers the use of web APIs and linked data principles. The hypothesis is that applications are increasingly accessing the multifarious and heterogeneous web of data. The proposal is in contrast to platform to develop applications consuming linked data in a declarative and modular way. It considers in detail the functional language – the platform gives access to, which is based on SPARQL (the standard query language for linked data) and on the dataflow paradigm. By adopting a declarative style, it favours the development of modules that can be reused in various execution contexts, thus saving on the repetitive effort.