

Editor's Notes

Welcome to the September 2014 issue of the ACM SIGMOD Record!

The issue opens with a Database Principles article by Deutsch, Hull, and Vianu on automatic verification of database-centric systems. This work is motivated by the observation that the emerging high-level specification tools for database-centric systems provide a natural target for verification. The results described by this article suggest that verification may indeed be feasible for significant classes of database-driven systems, specifically under the two representative models of business artifacts and database-driven web services where restrictions can be placed to guarantee decidability of verification. The article concludes by outlining several directions for future work, including both theoretical and practical aspects. This article introduces the reader to an important research topic, and also facilitates future research at the confluence of database and computer-aided verification areas.

The Research and Vision Articles Column features an article by Guo, Jensen, and Yang on data management issues in the transportation sector. The article is driven by the vision that continued proliferation of sensors and mobile devices, combined with the drive towards open data, will result in rapidly increasing volumes of transportation data. Furthermore, efficient and effective analysis of “big transportation data” will enable people to extract new, important transportation knowledge. Towards this goal, this article describes the transportation data, presents key challenges related to the extraction of thorough, timely, and trustworthy traffic knowledge to achieve total traffic awareness, and closes by outlining the services that will be enabled. Overall, this article provides a good introduction to big transportation data, with the aim to motivate more researchers to work on the related data management problems.

The Surveys Column features two articles. Braganholo and Mattoso provide a survey on XML fragmentation. This is a key issue in distributed (e.g., MapReduce) processing where large volumes of XML data are split into pieces and distributed to different compute nodes, and queries are split accordingly to exploit parallel processing. However, there is no consensus in the database literature on what an XML fragment is. This article surveys different XML fragmentation techniques, discussing their features and highlighting their limitations. Such discussion will enable the XML user to choose the appropriate technique based on the target query-processing scenario. The second article, by Santos, Bernardino, and Vieira, surveys Database Intrusion Detection Systems (DIDS). It presents a variety of existing intrusion detection techniques, and discusses how these IDIS are applied in each database context, pointing out their weaknesses given typical user workloads and characteristics of the environment. This discussion is followed by the identification of a set of challenges and opportunities, as well as requirements and guidelines to drive the development of new or improved DIDS.

The Systems and Prototypes Column features an integrated system for mining, querying, and managing Web document corpora, developed by Mousavi, Atzori, Gao, and Zaniolo. This system is designed for Wikipedia's InfoBoxes, which provide the main knowledge source for many applications on the Web while suffering from incompleteness, inconsistencies, and inaccuracies. To overcome these problems, the proposed system integrates the *IBminer* system that extracts InfoBox information by text-mining Wikipedia pages; the *IKBStore* system that integrates mined text with other sources to build a knowledge base; and a user-friendly interface that supports knowledge editing and query-by-example functions needed for managing a knowledge base.

In the Research Centers Column, Kemper and Neumann outline the research agenda of the database group at Technische Universität München (TUM). The research at TUM has a long-term goal of developing a high-performance database engine that unites OLTP and OLAP into a single system. For this purpose, a hybrid main memory database system, called HyPer, has been developed. HyPer supports both OLTP and

OLAP applications on the same database state by exploiting the OS/processor-support for virtual memory management. This article describes the key design choices and reviews recent work on a variety of topics, including NUMA-aware many-core parallelism, in-memory database compression, indexing on main-memory databases, and lock-free synchronization by hardware transactional memory. This line of research has led to a number of awards granted to the members of the TUM database team.

This issue features two event reports. First, it is our pleasure to include the Beckman Report on Database Research. Every few years a group of database researchers meet to discuss the state of database research, its impact on practice, and important new directions. This report summarizes the discussion and conclusions of the eighth such meeting, held October 14-15, 2013 in Irvine California. The meeting identified Big Data as a defining challenge of our time, and argued that enormous opportunities are presented to the database community to make transformative impact. In particular, this report discusses research challenges in five areas: scalable big/fast data infrastructures; coping with diversity in the data management landscape; end-to-end processing and understanding of data; cloud services; and managing the diverse roles of people in the data life cycle. The article also discusses community challenges, in particular, the need to rethink our approach to teaching data management technologies. The second article in the column, by Varde and Tatti, reports on the PhD Forum held at the IEEE International Conference on Data Mining (ICDM) 2013. The report surveys the papers from the forum, covering both core research and applied research in data mining, and discusses open issues that provide a scope for further research.

This issue closes with the call for papers for EuroSys 2015, to be held April 21-24, 2015 in Bordeaux, France, and the call for participation for SoCC, to be held November 3-5, 2014 in Seattle, Washington.

On behalf of the SIGMOD Record Editorial board, I hope that you will all enjoy reading the September 2014 issue of the SIGMOD Record.

Your submissions to the Record are welcome via the submission site:

<http://sigmod.hosting.acm.org/record>

Prior to submitting, you are encouraged to read the Editorial Policy on the SIGMOD Record's Web site (<http://www.sigmod.org/publications/sigmod-record/sigmod-record-editorial-policy>).

Yanlei Diao
September 2014

Past SIGMOD Record Editors:

Harrison R. Morse (1969)
Daniel O'Connell (1971 – 1973)
Randall Rustin (1974-1975)
Douglas S. Kerr (1976-1978)
Thomas J. Cook (1981 – 1983)
Jon D. Clark (1984 – 1985)
Margaret H. Dunham (1986 – 1988)
Arie Segev (1989 – 1995)
Jennifer Widom (1995 – 1996)
Michael Franklin (1996 – 2000)
Ling Liu (2000 – 2004)
Mario Nascimento (2005 – 2007)
Alexandros Labrinidis (2007 – 2009)
Ioana Manolescu (2009-2013)