Lecture Notes
in Business Information Processing 126

Series Editors

Malu Castellanos
Umeshwar Dayal
Wolfgang Lehner (Eds.)

Enabling Real-Time Business Intelligence

5th International Workshop, BIRTE 2011
Held at the 37th International Conference
on Very Large Databases, VLDB 2011
Seattle, WA, USA, September 2, 2011
Revised Selected Papers

Springer

Volume Editors

Malu Castellanos
Hewlett-Packard
Palo Alto, CA, USA
E-mail: malu.castellanos@hp.com

Umeshwar Dayal
Hewlett-Packard
Palo Alto, CA, USA
E-mail: umeshwar.dayal@hp.com

Wolfgang Lehner
Dresden University of Technology
Dresden, Germany
E-mail: wolfgang.lehner@tu-dresden.de

ISSN 1865-1348 e-ISSN 1865-1356
ISBN 978-3-642-33499-3 e-ISBN 978-3-642-33500-6
DOI 10.1007/978-3-642-33500-6
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2012947035

ACM Computing Classification (1998): H.3, J.1, H.2

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Business intelligence has evolved into a multi-billion dollar market over the last decade. Since the early years of data warehousing, business needs have constantly posed new requirements on state-of-the-art business intelligence systems. In today's competitive and highly dynamic environment, providing insight does not merely require analysis of the existing data. Deriving actionable intelligence demands the efficient processing of a vast amount of information, in order to arrive at a timely representation of the state of an enterprise as well as of emerging trends. Prediction models must be used in order to assist with the derivation of actions from the current state of the enterprise and the market, taking into account the uncertainty of the prediction. Moreover, the increasing use of twitter, blogs, and other media means that business intelligence cannot restrict itself to only dealing with structured information. More and more information sources of varying kind have to be integrated, starting with the vast amount of textual information in corporate intranets and the Web. However, because of media convergence, future business intelligence will also have to consider audio and video streams as further information sources. The end goal is to support better and timelier decision making, enabled by the availability of up-to-date, high-quality information.

Although there has been progress in this direction and many companies are introducing products toward meeting this goal, there is still a long way to go. In particular, the whole life cycle of business intelligence requires new techniques and methodologies capable of dealing with the new requirements imposed by the new generation of business intelligence applications. From the capturing of real-time business data to the transformation and delivery of actionable information, all the stages of the business intelligence cycle call for new algorithms and paradigms as the basis of new functionalities including business intelligence over text data, ensuring information quality, dealing with the uncertainty of prediction models, nested complex events, and optimizing complex ETL workflows, just to name a few.

The series of BIRTE workshops aims to provide a forum to discuss and advance the foundational science and engineering required to enable real-time business intelligence and the novel applications and solutions that build on these foundational techniques. Following the success of our previous workshops colocated with the VLDB conferences in Seoul, Auckland, Lyon, and Singapore, our fifth workshop was held in Seattle on September 2, 2011.

After the official opening of the workshop, Guy Lohman (IBM Almaden Research Center, USA) delivered an excellent and engaging keynote entitled "Blink: Not Your Father's Database," where he outlined the key concepts of the IBM Accelerator for analytical database queries. The first session on "Innovative System Architectures" included two research papers. The contribution by Qiming Chen

et al. (HP Labs, USA) outlined the core concepts of MemcacheSQL, a SQL Cache engine for massive scale-out query processing. The second paper presented by Florian Huebner with co-authors from Hasso Plattner Institute and SAP proposed a cost-aware strategy for merging differential stores in column-oriented in-memory DBMS. The session was closed with an invited talk delivered by Jose Blakeley from Microsoft Corporation on the architecture of the Microsoft SQL Server Parallel Data Warehouse. In this presentation Jose described how the system exploits parallelism to provide a foundation for real-time enabled enterprises. The second session on Novel Query Support included two research papers. "Relax and Let the Database do the Partitioning Online" by Alekh Jindal and Jens Dittrich from Saarland University, Germany and "Adaptive Processing of Multi-Criteria Decision Support Queries" by Venkatesh Raghavan et al. (Worcester Polytechnic Institute, USA). The final session on "Innovative Applications" again was a combination of an invited talk and the presentation of a research paper. The invited talk was delivered by Shilpa Lawande and Lakshmikant Shrinivas from Vertica, a Hewlett Packard Company on "Scalable Social-Graphing Analytics with the Vertica Analytic Platform." They gave an impressive demo on how suitable modern database platforms such as Vertica are for non-standard applications like graph processing on extremely large databases. The final presentation within the application part of the workshop was delivered by Amit Rustagi (eBay), and outlined the requirements and architecture of "A Near Real-Time Personalization for eCommerce Platform."

We wish to express special thanks to the Program Committee members for helping us prepare an interesting program. To our keynote speakers, presenters, and attendees, we express our appreciation for sharing their work and the lively discussions that made this workshop a great forum for exchanging new ideas. We thank the VLDB 2011 organizers for their help and organizational support. Finally, we would like to thank Maik Thiele, our Publication Chair.

September 2012

Malu Castellanos
Umeshwar Dayal
Wolfang Lehner

Organization

Organizing Committee

General Chair

Umeshwar Dayal	Hewlett-Packard, USA

Program Committee Chairs

Malu Castellanos	Hewlett-Packard, USA
Wolfgang Lehner	TU Dresden, Germany

Program Committee

Christof Bornhövd	SAP Labs, USA
Ben Chin Ooi	National University of Singapore, Singapore
Chetan Gupta	HP Labs, USA
Howard Ho	IBM, USA
Alfons Kemper	TUM, Germany
Christian Konig	Microsoft, USA
Alexander Loeser	University of Technology Berlin, Germany
Jose Norberto Mazon	University of Alicante, Spain
Renee Miller	University of Toronto, Canada
Torben B. Pedersenv	Aalborg University, Denmark
Elke Rundensteiner	WPI, USA
Donovan Schneider	SalesForce, USA
Eric Simon	SAO-BO, France
Nesime Tatbul	ETH Zurich, Switzerland
Florian Waas	Greenplum, USA

Publication Chair

Maik Thiele	Technische Universität Dresden, Germny

Table of Contents