STREAM DATA MANAGEMENT

ADVANCES IN DATABASE SYSTEMS

Series Editor

Ahmed K. Elmagarmid

Purdue University
West Lafayette, IN 47907

Other books in the Series:

FUZZY DATABASE MODELING WITH XML, *Zongmin Ma,* ISBN 0-387-24248-1; e-ISBN 0-387-24249-X

MINING SEQUENTIAL PATTERNS FROM LARGE DATA SETS, *Wei Wang and Jiong Yang;* ISBN 0-387-24246-5; e-ISBN 0-387-24247-3

ADVANCED SIGNATURE INDEXING FOR MULTIMEDIA AND WEB APPLICATIONS, *Yannis Manolopoulos, Alexandros Nanopoulos, Eleni Tousidou;* ISBN: 1-4020-7425-5

ADVANCES IN DIGITAL GOVERNMENT, *Technology, Human Factors, and Policy,* *edited by William J. McIver, Jr. and Ahmed K. Elmagarmid;* ISBN: 1-4020-7067-5

INFORMATION AND DATABASE QUALITY, *Mario Piattini, Coral Calero and Marcela Genero;* ISBN: 0-7923- 7599-8

DATA QUALITY, *Richard Y. Wang, Mostapha Ziad, Yang W. Lee:* ISBN: 0-7923-7215-8

THE FRACTAL STRUCTURE OF DATA REFERENCE: *Applications to the Memory Hierarchy,* *Bruce McNutt;* ISBN: 0-7923-7945-4

SEMANTIC MODELS FOR MULTIMEDIA DATABASE SEARCHING AND BROWSING, *Shu-Ching Chen, R.L. Kashyap, and Arif Ghafoor;* ISBN: 0-7923-7888-1

INFORMATION BROKERING ACROSS HETEROGENEOUS DIGITAL DATA: A Metadata-based Approach, *Vipul Kashyap, Amit Sheth;* ISBN: 0-7923-7883-0

DATA DISSEMINATION IN WIRELESS COMPUTING ENVIRONMENTS, *Kian-Lee Tan and Beng Chin Ooi;* ISBN: 0-7923-7866-0

MIDDLEWARE NETWORKS: Concept, Design and Deployment of Internet Infrastructure, *Michah Lerner, George Vanecek, Nino Vidovic, Dad Vrsalovic;* ISBN: 0-7923-7840-7

ADVANCED DATABASE INDEXING, *Yannis Manolopoulos, Yannis Theodoridis, Vassilis J. Tsotras;* ISBN: 0-7923-7716-8

MULTILEVEL SECURE TRANSACTION PROCESSING, *Vijay Atluri, Sushil Jajodia, Binto George* ISBN: 0-7923-7702-8

FUZZY LOGIC IN DATA MODELING, *Guoqing Chen* ISBN: 0-7923-8253-6

INTERCONNECTING HETEROGENEOUS INFORMATION SYSTEMS, *Athman Bouguettaya, Boualem Benatallah, Ahmed Elmagarmid* ISBN: 0-7923-8216-1

FOUNDATIONS OF KNOWLEDGE SYSTEMS: With Applications to Databases and Agents, *Gerd Wagner* ISBN: 0-7923-8212-9

DATABASE RECOVERY, *Vijay Kumar, Sang H. Son* ISBN: 0-7923-8192-0

For a complete listing of books in this series, go to http://www.springeronline.com

STREAM DATA MANAGEMENT

edited by

Nauman A. Chaudhry
University of New Orleans, USA

Kevin Shaw
Naval Research Lab, USA

Mahdi Abdelguerfi
University of New Orleans, USA

Springer

Nauman A. Chaudhry
University of New Orleans
USA

Kevin Shaw
Naval Research Lab
USA

Mahdi Abdelguerfi
University of New Orleans
USA

Library of Congress Cataloging-in-Publication Data

A C.I.P. Catalogue record for this book is available from the
Library of Congress.

STREAM DATA MANAGEMENT
edited by
Nauman A. Chaudhry
Kevin Shaw
Mahdi Abdelguerfi

Advances in Database Systems
Volume 30

ISBN 0-387-24393-3 e-ISBN 0-387-25229-0

Cover by Will Ladd, NRL Mapping, Charting and Geodesy Branch
utilizing NRL's GIDB® Portal System that can be utilized at
http://dmap.nrlssc.navy.mil

Printed on acid-free paper.

Printed in the United States of America.

9 8 7 6 5 4 3 2 1 SPIN 11054597, 11403999

springeronline.com

Contents

List of Figures

List of Tables

Preface

In recent years, a new class of applications has emerged that requires managing data streams, i.e., data composed of continuous, real-time sequence of items. However, database management systems were originally developed to support business applications. The data in such applications is changed as a result of human-initiated transactions. Similarly data is queried as a result of human-initiated queries. The database management system acts as a passive repository for the data, executing the queries and transactions when these are submitted. However, this model of a database management system as a repository of relatively static data that is queried as a result of human interaction, does not meet the challenges posed by streaming applications.

A data stream is a possibly unbounded sequence of data items. Streaming applications have gained prominence due to both technical and business reasons. Technologically data is now available from a wide variety of monitoring devices, including sensors that are extremely cheap. Data from such devices is potentially unbounded and needs to be processed in real-time. Additionally businesses and Federal agencies now increasingly want to perform analysis on data much sooner than is possible with the current model of storing data in a data warehouse and performing the analysis off-line. Application domains requiring data stream management include military, homeland security, sensor networks, financial applications, network management, web site performance tracking, real-time credit card fraud detection, etc.

Streaming applications pose new and interesting challenges for data management systems. Such application domains require queries to be evaluated continuously as opposed to the one time evaluation of a query for traditional applications. Streaming data sets grow continuously and queries must be evaluated on such unbounded data sets. The monitoring aspect of many streaming applications requires support for reactive capabilities in real-time from data management systems. These, as well as other challenges, require a major rethink of almost all aspects of traditional database management systems to support streaming applications. Consequently, stream data management has been a very active area of research over the past few years.

The goal of this edited manuscript is to gather a coherent body of work spanning various aspects of stream data management. The manuscript comprises eight invited chapters by researchers active in stream data management. The collected chapters provide exposition of algorithms, languages, as well as systems proposed and implemented for managing streaming data. We expect this book will appeal to researchers already involved in stream data management, as well as to those starting out in this exciting and growing area.