

*Commenced Publication in 1973*

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

## Editorial Board

David Hutchison

*Lancaster University, UK*

Takeo Kanade

*Carnegie Mellon University, Pittsburgh, PA, USA*

Josef Kittler

*University of Surrey, Guildford, UK*

Jon M. Kleinberg

*Cornell University, Ithaca, NY, USA*

Friedemann Mattern

*ETH Zurich, Switzerland*

John C. Mitchell

*Stanford University, CA, USA*

Moni Naor

*Weizmann Institute of Science, Rehovot, Israel*

Oscar Nierstrasz

*University of Bern, Switzerland*

C. Pandu Rangan

*Indian Institute of Technology, Madras, India*

Bernhard Steffen

*TU Dortmund University, Germany*

Madhu Sudan

*Microsoft Research, Cambridge, MA, USA*

Demetri Terzopoulos

*University of California, Los Angeles, CA, USA*

Doug Tygar

*University of California, Berkeley, CA, USA*

Moshe Y. Vardi

*Rice University, Houston, TX, USA*

Gerhard Weikum

*Max Planck Institute for Informatics, Saarbruecken, Germany*

Abdelkader Hameurlain Josef Küng  
Roland Wagner Bernd Amann  
Philippe Lamarre (Eds.)

# Transactions on Large-Scale Data- and Knowledge- Centered Systems XI

Special Issue on Advanced Data Stream Management  
and Continuous Query Processing

## Editors-in-Chief

Abdelkader Hameurlain  
Paul Sabatier University, IRIT, Toulouse, France  
E-mail: hameur@irit.fr

Josef Küng  
Roland Wagner  
University of Linz, FAW, Austria  
E-mail: {jkueng, rrwagner}@faw.at

## Guest Editors

Bernd Amann  
Université Pierre et Marie Curie, Paris, France  
E-mail: bernd.amann@lip6.fr

Philippe Lamarre  
Université de Lyon, France  
E-mail: philippe.lamarre@liris.cnrs.fr

ISSN 0302-9743 (LNCS)	e-ISSN 1611-3349 (LNCS)
ISSN 1869-1994 (TLDKS)	
ISBN 978-3-642-45268-0	e-ISBN 978-3-642-45269-7
DOI 10.1007/978-3-642-45269-7	
Springer Heidelberg New York Dordrecht London	

Library of Congress Control Number: Applied for

CR Subject Classification (1998): H.2.8, H.2, H.3, F.1, C.2, J.1

© Springer-Verlag Berlin Heidelberg 2013

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

*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

An increasing number of applications continuously produce and consume high volumes of streaming data. Most of these applications are still developed from scratch and are rapidly overwhelmed by complex data management issues generated by the fast growth in the number of users, data items, and processing tasks. Data Stream Management Systems follow traditional database principles to facilitate the development of this kind of application via high-level programming interfaces and efficient data stream processing technologies built on declarative data stream models and languages, advanced query optimization algorithms, and new stream processing architectures.

This TLDKS Special Issue presents a representative selection of articles covering a wide range of important topics in advanced data stream management and processing of continuous queries. The book starts with application and development oriented contributions and ends with more theoretical results concerning data stream mining and analysis.

The first article “ASSIST: Access Controlled Ship Identification Stream” by Jianneng Cao et al. illustrates the application of algebraic rewriting techniques to avoid unauthorized access to streaming data. This article is particularly interesting because it also includes a real data stream management scenario involving maritime traffic control data of the port of Singapore shared among ship’s captains, port authorities, shipping and insurance companies, cargo owners, and other stakeholders.

The design and development of applications mixing traditional databases and streaming data is a complex task for which it is crucial to have suitable tools. In their article “The HIT Model: Workflow-Aware Event Stream Monitoring” Olga Poppe et al. propose a workflow model combining continuous queries and timed automata to model complex real time reactive behavior of stream processing applications.

In a different direction, but still focussing on application development, the article “P-Bench: Benchmarking in Data-Centric Pervasive Application Development” by Sabina Surdu et al. characterizes, evaluates, and compares different event-centric and data-centric platforms for the development of applications in pervasive environments. Their work is a first step towards a benchmark for data stream management systems.

The amount of information published and disseminated in the form of data streams is constantly increasing and generates new challenges in the data analysis and knowledge discovery domain. In their article “Incremental Mining of Lag Patterns in Time Series Databases”, Dhaval Patel et al. propose an adaptive and incremental approach to capture localized repeated associative relationships from multiple time series. In spite of the interest of such methods, for example for financial analysis, their use is currently limited by an excessive consumption

of resources. The incremental nature of their approach opens up the possibility a better scalability. Finally, Emmanuelle Anceaume et al. focussed their work entitled “On the Power of the Adversary to Solve the Node Sampling Problem” on the sampling of node identifier streams in unstructured peer-to-peer systems with malicious nodes. Node identifier streams constitute the keystone for many services (information dissemination, load balancing, data caching, etc.) and malicious nodes try to bias these streams, for example, to isolate some participant. The authors formally study the problem of obtaining uniform and fresh sampling strategies over input data streams generated by large open systems containing malicious participants.

The publication of this TLDKS Special Issue would not have been possible without the help of many people. First, we would like to thank all the authors who have submitted their articles and all the reviewers for their valuable efforts to ensure the high quality of this special edition. We also thank Abdelkader Hameurlain, Josef Küng, and Roland Wagner, Editors-in-Chief of the TLDKS journal, for their confidence and help. Finally, we are particularly grateful to Gabriela Wagner for her assistance in the preparation of these proceedings.

October 2013

Bernd Amann  
Philippe Lamarre

## Editorial Board

Reza Akbarinia  
Stéphane Bressan  
Francesco Buccafurri

Yuhan Cai  
Qiming Chen  
Tommaso Di Noia  
Dirk Draheim  
Johann Eder  
Stefan Fenz  
Georg Gottlob  
Anastasios Gounaris  
Theo Härder

Dieter Kranzlmüller

Philippe Lamarre  
Lenka Lhotská  
Vladimir Marik  
Mukesh Mohania  
Tetsuya Murai  
Gultekin Ozsoyoglu  
Torben Bach Pedersen  
Günther Pernul  
Klaus-Dieter Schewe  
Makoto Takizawa  
David Taniar  
A Min Tjoa  
Chao Wang

INRIA, France  
National University of Singapore, Singapore  
Università Mediterranea di Reggio Calabria,  
Italy  
A9.com, USA  
HP-Lab, USA  
Politecnico di Bari, Italy  
University of Innsbruck, Austria  
Alpen Adria University Klagenfurt, Austria  
Vienna University of Technology, Austria  
Oxford University, UK  
Aristotle University of Thessaloniki, Greece  
Technical University of Kaiserslautern,  
Germany  
Ludwig-Maximilians-Universität München,  
Germany  
University of Nantes, France  
Technical University of Prague, Czech Republic  
Technical University of Prague, Czech Republic  
IBM India, India  
Hokkaido University, Japan  
Case Western Reserve University, USA  
Aalborg University, Denmark  
University of Regensburg, Germany  
University of Linz, Austria  
Seikei University Tokyo, Japan  
Monash University, Australia  
Vienna University of Technology, Austria  
Oak Ridge National Laboratory, USA

## Reviewers

Reza Akbarinia  
Bernd Amann  
Stephane Bressan  
Peter M. Fischer  
Pedro Furtado  
Joao Gama  
Lukasz Golab

INRIA, Nantes, France  
LIP6 – UPMC, France  
National University of Singapore, Singapore  
University of Freiburg, Germany  
Universidade de Coimbra, Portugal  
University of Porto, Portugal  
University of Waterloo, Canada

Georges Hebrail  
Cyril Labbe  
Philippe Lamarre  
Sebastian Michel  
Jean-Marc Petit  
Evaggelia Pitoura  
Bernhard Seeger

EDF R&D, France  
Grenoble University/LIG, France  
LIRIS - INSA Lyon, France  
Max-Planck-Institut, Germany  
LIRIS, INSA Lyon, France  
University of Ioannina, Greece  
Philipps-University Marburg, Germany

# Table of Contents

ASSIST: Access Controlled Ship Identification Streams . . . . .	1
<i>Jianneng Cao, Thomas Kister, Shili Xiang, Baljeet Malhotra, Wee-Juan Tan, Kian-Lee Tan, and Stéphane Bressan</i>	
The HIT Model: Workflow-Aware Event Stream Monitoring . . . . .	26
<i>Olga Poppe, Sandro Giessel, Elke A. Rundensteiner, and François Bry</i>	
P-Bench: Benchmarking in Data-Centric Pervasive Application Development . . . . .	51
<i>Sabina Surdu, Yann Gripay, Vasile-Marian Scuturici, and Jean-Marc Petit</i>	
Efficient Mining of Lag Patterns in Evolving Time Series . . . . .	76
<i>Dhaval Patel, Wynne Hsu, and Mong Li Lee</i>	
On the Power of the Adversary to Solve the Node Sampling Problem . . .	102
<i>Emmanuelle Anceaume, Yann Busnel, and Sébastien Gambs</i>	
<b>Author Index</b> . . . . .	127