Lecture Notes in Computer Science 8381

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Aastha Madaan   Shinji Kikuchi
Subhash Bhalla (Eds.)

Databases in Networked Information Systems

9th International Workshop, DNIS 2014
Aizu-Wakamatsu, Japan, March 24-26, 2014
Proceedings

Springer

Volume Editors

Aastha Madaan
Shinji Kikuchi
Subhash Bhalla

University of Aizu
Graduate Department of Computer and Information Systems
Ikki Machi, Aizu-Wakamatsu, Fukushima 965-8580, Japan

Email:
d8131102@u-aizu.ac.jp
shinji.kikuchi@kne.biglobe.ne.jp
bhalla@u-aizu.ac.jp

ISSN 0302-9743
e-ISSN 1611-3349
ISBN 978-3-319-05692-0
e-ISBN 978-3-319-05693-7
DOI 10.1007/978-3-319-05693-7
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014933394

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

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 data analytics in astronomy and sciences depends on computing infrastructure. Such scientific exploration is beneficial for large-scale public utility services, either directly or indirectly. Many research efforts are being made in diverse areas, such as big data analytics and cloud computing, sensor networks, and high-level user interfaces for information accesses by users. Government agencies in many countries plan to launch facilities in education, health-care, and information support as a part of an e-government initiative. In this context, information interchange management has become an active research field. A number of new opportunities have evolved in design and modeling 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.

The 9th International Workshop on Databases in Networked Information Systems (DNIS) was held during March 24–26, 2014, at the University of Aizu in Japan. The workshop program included research contributions and invited contributions. A view of the research activity in information interchange management and related research issues was provided by the sessions on related topics. The keynote address has been contributed by - Prof. Divyakant Agrawal. The section on "Astronomical Data Management" has an invited contribution from Dr. Florin Rusu. The following section on "Business Data Analytics and Visualization," has an invited contribution from Prof. Marcin Paprzycki. The section on "Business Data Analytics in Sciences," includes the invited contributions of Dr. Lukas Pichl. The section on "Business Data Analytics in Astronomy," has an invited contribution by Prof. Thomas A. Prince. I would like to thank the members of the Program Committee for their support and all authors who considered DNIS 2014 for presenting research contributions.

The sponsoring organizations and the Steering Committee deserve praise for the support they provided. A number of individuals contributed to the success of the workshop. I thank Dr. Umeshwar Dayal, Prof. J. Biskup, Prof. D. Agrawal, Dr. Cyrus Shahabi, Prof. T. Nishida, and Prof. Shrinivas Kulkarni for providing continuous support and encouragement.

The workshop received invaluable support from the University of Aizu. In this context, I thank Prof. Shigeaki Tsunoyama, President of the University of Aizu. Many thanks are also extended to the faculty members at the university for their cooperation and support.

March 2014

A. Madaan
S. Kikuchi
S. Bhalla

Organization

DNIS 2014 was organized by the Graduate Department of Information Technology and Project Management, University of Aizu, Aizu-Wakamatsu, Fukushima, PO 965-8580, (JAPAN).

Steering Committee

Divyakant Agrawal	University of California, USA
Hosagrahar V. Jagadish	University of Michigan, USA
Umeshwar Dayal	Hewlett-Packard Laboratories, USA
Masaru Kitsuregawa	University of Tokyo, Japan
Toyoaki Nishida	Kyoto University, Japan
Krithi Ramamritham	Indian Institute of Technology, Bombay, India
Cyrus Shahabi	University of Southern California, USA

Executive Chair

N. Bianchi-Berthouze	University College London, UK

Program Chair

S. Bhalla	University of Aizu, Japan

Publicity Committee Chair

Shinji Kikuchi	University of Aizu, Japan

Publications Committee Chair

Aastha Madaan	University of Aizu, Japan

Program Committee

D. Agrawal	University of California, USA
V. Bhatnagar	University of Delhi, India
P. Bottoni	University La Sapienza of Rome, Italy
L. Capretz	University of Western Untario, Canada
Richard Chbeir	Bourgogne University, France
G. Cong	Nanyang Technological University, Singapore

U. Dayal	Hewlett-Packard Laboratories, USA
Pratul Dublish	Microsoft Research, USA
Fernando Ferri	IRPPS - CNR, Rome, Italy
W.I. Grosky	University of Michigan-Dearborn, USA
J. Herder	University of Applied Sciences, Fachhochschule Düsseldorf, Germany
H.V. Jagadish	University of Michigan, USA
Sushil Jajodia	George Mason University, USA
Q. Jin	Waseda University, Japan
A. Kumar	Pennsylvania State University, USA
A. Mondal	Xerox Research, Bangaloru, India
K. Myszkowski	Max-Planck-Institut für Informatik, Germany
Alexander Pasko	Bournemouth University, UK
L. Pichl	International Christian University, Tokyo, Japan
P.K. Reddy	International Institute of Information Technology, Hyderabad, India
C. Shahabi	University of Southern California, USA
M. Sifer	Sydney University, Australia
F. Wang	Microsoft Research, USA

Sponsoring Institution

Center for Strategy of International Programs, University of Aizu
Aizu-Wakamatsu City, Fukushima P.O. 965-8580, Japan.

Abstracts

Data Exploration in Large Area Time-Domain Sky Surveys: Current Practice and Future Needs

Thomas A. Prince

Jet Propulsion Laboratory
California Institute of Technology
MS 169-327, 4800 Oak Grove Dr., Pasadena, CA 91109
prince@srl.caltech.edu

Abstract. Time-domain astronomical surveys are playing an increasingly important role in astronomy, driven by the availability of large format digital sensors and powerful computational resources. Examples of such surveys include the Palomar Transient Factory (PTF), the Catalina Real-time Transient Survey (CRTS), the PanSTARRS survey, and the future Large Synoptic Survey Telescope (LSST) and Zwicky Transient Facility (ZTF). These surveys typically are sensitive to on order of a billion or more objects over the course of a year and the challenge is to find the few hundred to a few thousand most interesting variable sources among the total collection of objects. These include supernovae, gamma-ray bursts, near-earth asteroids, cataclysmic variables, and a large number of other exotic objects. While technical challenges exist in the area of database management and storage, the most difficult challenges are algorithmic. We will discuss some of the principal challenges, including the unique problems associated with the intrinsically heterogeneous collection of time-domain data and the application of machine learning techniques to automatic identification of interesting variable sources.

Towards an Intelligent Astronomical Event Broker: Automated Transient Classification and Follow-Up Optimization

Przemek Wozniak, Ph.D.,
Optical Sensing Team Leader

Space and Remote Sensing, ISR-2, MS-B244
Los Alamos National Laboratory, Los Alamos, NM 87545
wozniak@lanl.gov

Abstract. In order to succeed, the massive time-domain surveys of the future must automatically identify actionable information from the torrent of imaging data, classify emerging events, and optimize the follow-up strategy. The impedance mismatch between the high rate of transient detection and limited follow-up resources is growing rapidly. Despite a pressing need to delegate increasingly complex data fusion and inference tasks to machines, the available technology toolbox remains underutilized in astronomy. To address these challenges, we are developing an autonomous, distributed event broker that will integrate cutting edge machine learning algorithms with high performance computing infrastructure. The talk will give an overview of various efforts in this area, including recent progress on image level variability detection, spectral classification using low resolution spectra and dynamic coalition management approach to follow-up.

Machine-Learning Enabled Stellar Classification and the Prediction of Fundamental Atmospheric Parameters From Photometric Light Curves

Adam A. Miller

Jet Propulsion Laboratory
California Institute of Technology
MS 169-327, 4800 Oak Grove Dr., Pasadena, CA 91109
amiller@astro.caltech.edu

Abstract. The falling costs of computing and CCD detectors has led to a great boom in wide-field time-domain surveys during the past decade, with several new surveys expected prior to the arrival of the Large Synoptic Survey Telescope (LSST). This observational boon, however, comes with a catch: the data rates from these surveys are so large that discovery techniques heavily dependent on human intervention are becoming unviable. In this talk I will detail new methods, which utilize semi-supervised machine-learning algorithms, to automatically classify the light curves of time-variable sources. Using these methods, we have produced a data-driven probabilistic catalog of variables found in the All Sky Automated Survey (ASAS). I will also present a new machine-learning-based framework for the prediction of the fundamental stellar parameters, Teff, log g, and [Fe/H], based on the photometric light curves of variable stellar sources. The method was developed following a systematic spectroscopic survey of stellar variability. I will demonstrate that, for variable sources, the machine-learning model can determine Teff, log g, and [Fe/H] with a typical scatter of 130 K, 0.38 dex, and 0.26 dex, respectively, without obtaining a spectrum. Instead, the random-forest-regression model uses SDSS color information and light-curve features to infer stellar properties. The precision of this method is competitive with what can be achieved with low-resolution spectra. These results are an important step on the path to the efficient and optimal extraction of information from future time-domain experiments, such as LSST. We argue that this machine-learning framework, for which we outline future possible improvements, will enable the construction of the most detailed maps of the Milky Way ever created.

Astrophysical Image Modeling

Robert M. Quimby

Kavli Institute for the Physics and Mathematics of the Universe (WPI)
Todai Institutes for Advanced Study
the University of Tokyo
5-1-5 Kashiwa-no-Ha, Kashiwa City, Chiba 277-8583, Japan
robert.quimby@ipmu.jp

Abstract. Discovery time variable objects in astronomical data requires that the same patch of sky be observed at least twice and preferably more often. The first observation can be used to predict what that same patch of sky will look like in subsequent observations to finite accuracy (given noise and changes in the atmosphere or instrument through which the observations are made). In other words, a model can be constructed by some transformation of the data obtained at the first epoch to match the observational characteristics of a second observation. Comparing this model to the actual observations, candidates for time variability will stand out as objects that are poorly predicted by the model. Here I discuss how these models of the astrophysical sky are currently constructed using the Palomar Transient Factory as a specific example. I will point out some of the weak points of this process and discuss some new techniques which may prove beneficial for the discovery of time variable objects.

Table of Contents

Business Data Analytics in Sciences

Geo-spatial Decision Making and Query Languages