Preface

For over two decades, the International Conference on Management of Data (COMAD), modeled along the lines of ACM SIGMOD, has been the premier international database conference hosted in India by Division II of Computer Society of India, CSI. The first COMAD was held in Hyderabad in 1989, and it is wonderful that in its 25th year it has returned to Hyderabad. The 20th edition in the COMAD series is held at the campus of International Institute of Information Technology (IIIT) Hyderabad, from December 17-19, 2014.

COMAD seeks to provide the community of researchers, practitioners, developers and users of data management technologies, a forum to present and discuss problems, solutions, innovations, experiences and emerging trends. Keeping with the fast changing landscape of data management and analytics, the scope of COMAD 2014 has evolved to include emerging topics in Big Data Analytics, Web, Information Retrieval, Data Mining and Machine Learning in addition to the traditional topics in data management.

This year's call for papers attracted 63 research submissions from across the world. Each research paper was rigorously reviewed by at least three members of the program committee, which featured 26 data management experts from academia and industry from 4 different continents. After in-depth discussions, we selected 6 high-quality research papers for presentation at the conference, 2 industry research papers, 6 poster presentations and 3 demonstrations.

COMAD 2014 features three keynote talks by Prof. S. Muthukrishnan (Rutgers University and Microsoft Research), Prof. Renée Miller (University of Toronto, Canada), and Srini V. Srinivasan (Founder and VP of Engg. and Operations, Aerospike Inc.). The keynotes focus on very different aspects of "Big-Data" challenge - algorithms, curation and engineering. The program also hosts 3 tutorials from leading experts covering entity extraction and disambiguation, data mining over large-scale software repositories and how it can help software engineering, and mining massive-scale web repositories. We also continued the tradition started by COMAD in 2010 to invite Indian authors of papers published in premier international conferences to present their work at COMAD. This year features 2 papers from SIGMOD, and one paper each from PVLDB, KDD and ICDE from this year.

This time, COMAD 2014, also has the opportunity to have a special invited session with Prof. Jayant Haritsa (IISc) who was awarded the prestigious Infosys Prize this year, adding to an already long list of his honors.

To ensure visibility of COMAD beyond this conference, these proceedings will also be available through ACM SIGMOD and DBLP.

We would like to thank all the members of the COMAD Organizing Committee and the COMAD Program Committee for their generous support, enabling us to put together such a high-quality program. We are also grateful for the support and generosity of our sponsors. Without our silver sponsors Microsoft, Google, Infosys and Honeywell as well as our bronze sponsor Progress, this conference would not be possible. We also thank IIIT-Hyderabad for providing a campus for the conference. Finally, we acknowledge the sustained cooperation and assistance extended by the Computer Society of India in organizing this event.

In closing, we welcome you to the COMAD 2014 conference in Hyderabad and hope you will have a fruitful and stimulating experience.

Kamal Karlapalem

IIIT-Hyderabad, Hyderabad, India

(General Chair)

Divesh Srivastava

AT&T Labs Research, USA

Srikanta Bedathur

IBM Research, India

(Program Co-Chairs)

Satyanarayana R Valluri

École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

(Proceedings Chair)

Organizing Committee

GENERAL CHAIR	Kamal Karlapalem, IIIT Hyderabad
PROGRAM CHAIRS	Srikanta Bedathur, IBM Research, India Divesh Srivastava, AT&T Labs-Research
INDUSTRY CHAIR	Srinivasan Seshadri, Zettata
TUTORIALS & PANELS CHAIR	Sameep Mehta, IBM Research, India
POSTER & DEMO CHAIR	Manish Gupta, Microsoft India
PROGRAMMING CHALLENGE CHAIR	Arnab Bhattacharya, IIT Kanpur
WEB & PROCEEDINGS CHAIR	Satyanarayana R Valluri, EPFL, Switzerland
LOCAL ARRANGEMENTS CHAIR	P. Radhakrishna, Infosys, India

Program Committee

Srikanta Bedathur	IBM Research
Arnab Bhattacharya	Indian Institute of Technology, Kanpur
Indrajit Bhattacharya	IBM India Research Lab
Gautam Das	University of Texas at Arlington, USA
Mahashweta Das	HP Labs, Palo Alto
Prasad Deshpande	IBM Research - India
Lipika Dey	TCS Innovation Lab Delhi
Niloy Ganguly	Indian Institute of Technology Kharagpur
Vikram Goyal	IIIT-Delhi
Manish Gupta	Microsoft
Jayant Haritsa	Indian Institute of Science, Bangalore
Katja Hose	Aalborg University
Kalapriya Kannan	IBM
Gjergji Kasneci	Hasso-Plattner-Institute
Sameep Mehta	IBM Research
Karin Murthy	IBM Research
Aditya Parameshwaran	UIUC
Dhaval Patel	IIT Ropar
Krishna Reddy Polepalli	IIIT-H
Vikram Pudi	IIIT-H
Maya Ramanath	IIT Delhi
Sayan Ranu	IIT Madras
Ralf Schenkel	Universitaet Passau
Seshadri Srinivasan	Zettata
Divesh Srivastava	AT&T Labs-Research
S. Sudarshan	IIT Bombay

Table of Contents

Keynotes

Tutorials

Research Papers

Industry Papers

Poster Presentations

Demos