Advances in Intelligent Systems and Computing 554

Sanjiv K. Bhatia
Krishn K. Mishra
Shailesh Tiwari
Vivek Kumar Singh *Editors*

Advances in Computer and Computational Sciences

Proceedings of ICCCCS 2016, Volume 2

Springer

Advances in Intelligent Systems and Computing

Volume 554

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

About this Series

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing.

The publications within "Advances in Intelligent Systems and Computing" are primarily textbooks and proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

Advisory Board

More information about this series at http://www.springer.com/series/11156

Sanjiv K. Bhatia · Krishn K. Mishra
Shailesh Tiwari · Vivek Kumar Singh
Editors

Advances in Computer and Computational Sciences

Proceedings of ICCCCS 2016, Volume 2

Springer

Editors
Sanjiv K. Bhatia
Department of Mathematics and Computer Science
University of Missouri
St. Louis, MO
USA

Krishn K. Mishra
Department of Computer Science and Engineering
Motilal Nehru National Institute of Technology
Allahabad, Uttar Pradesh
India

Shailesh Tiwari
CSED
ABES Engineering College
Ghaziabad, Uttar Pradesh
India

Vivek Kumar Singh
Department of Computer Science
Banaras Hindu University
Varanasi, Uttar Pradesh
India

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-981-10-3772-6 ISBN 978-981-10-3773-3 (eBook)
https://doi.org/10.1007/978-981-10-3773-3

Library of Congress Control Number: 2017931526

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

The ICCCCS is a major multidisciplinary conference organized with the objective of bringing together researchers, developers and practitioners from academia and industry working in all areas of computer and computational sciences. It is organized specifically to help computer industry to derive the advances of next generation computer and communication technology. Researchers are invited to present the latest developments and technical solutions.

Technological developments all over the world are dependent upon globalization of various research activities. Exchange of information and innovative ideas are necessary to accelerate the development of technology. Keeping this ideology in preference, Aryabhatta College of Engineering & Research Center, Ajmer, India, has come up with an event—International Conference on Computer, Communication and Computational Sciences (ICCCCS-2016) during August 12–13, 2016.

Ajmer, situated in the heart of India, just over 130 km southwest of Jaipur, is a burgeoning town on the shore of the Ana Sagar Lake, flanked by barren hills. Ajmer has historical strategic importance and was ransacked by Mohammed Gauri on one of his periodic forays from Afghanistan. Later, it became a favorite residence of the mighty Mughals. The city was handed over to the British in 1818, becoming one of the few places in Rajasthan controlled directly by the British rather than being part of a princely state. The British chose Ajmer as the site for Mayo College, a prestigious school opened in 1875 exclusively for the Indian Princes, but today open to all those who can afford the fees. Ajmer is a perfect place that can be symbolized for demonstration of Indian culture and ethics and display of perfect blend of a plethora of diverse religions, communities, cultures, linguistics, etc., all coexisting and flourishing in peace and harmony. This city is known for the famous Dargah Sharif, Pushkar Lake, Brahma Temple, and many more evidences of history.

This is for the first time Aryabhatta College of Engineering & Research Center, Ajmer, India, is organizing International Conference on Computer, Communication and Computational Sciences (ICCCCS 2016), with a foreseen objective of enhancing the research activities at a large scale. Technical Program Committee and

Advisory Board of ICCCCS include eminent academicians, researchers, and practitioners from abroad as well as from all over the nation.

In this volume, the selected manuscripts have been subdivided into various tracks named 'Intelligent Hardware and Software Design', 'Advanced Communications', 'Power and Energy Optimization', 'Intelligent Image Processing', Advanced Software Engineering', 'IoT', 'ADBMS & Security', and 'Evolutionary and Soft Computing'. A sincere effort has been made to make it an immense source of knowledge for all, including 140 manuscripts in this collection. The selected manuscripts have gone through a rigorous review process and are revised by authors after incorporating the suggestions of the reviewers.

ICCCCS 2016 received 429 submissions from around 729 authors of 12 different countries such as USA, Iceland, China, Saudi Arabia, South Africa, Taiwan, and Malaysia. Each submission has been checked with anti-plagiarism software. On the basis of plagiarism report, each submission was rigorously reviewed by at least two reviewers with an average of 2.45 per reviewer. Even some submissions have more than two reviews. On the basis of these reviews, 140 high-quality papers were selected for publication in this proceedings volume, with an acceptance rate of 32.6%.

We are thankful to the speakers, delegates, and authors for their participation and their interest in ICCCCS as a platform to share their ideas and innovation. We are also thankful to the Prof. Dr. Janusz Kacprzyk, Series Editor, AISC, Springer and Mr. Aninda Bose, Senior Editor, Hard Sciences, Springer for providing continuous guidance and support. Also, we extend our heartfelt gratitude to the reviewers and Technical Program Committee Members for their concern and efforts in the review process. We are indeed thankful to everyone directly or indirectly associated with the conference organizing team leading it towards success.

Although utmost care has been taken in compilation and editing, however, a few errors may still occur. We request the participants to bear with such errors and lapses (if any). We wish you all the best.

Organizing Committee
ICCCCS 2016

Organizing Committee

General Chair
Dr. Amit Shastri, Chairman, Aryabhatta Academic Society, Ajmer, India

Program Chairs
Dr. K.K. Mishra, Motilal Nehru National Institute of Technology Allahabad, India
Dr. Munesh C. Trivedi, ABES, Engineering College, Ghaziabad, India

Conference Chair
Dr. Shailesh Tiwari, ABES, Engineering College, Ghaziabad, India

Conference Co-Chair
Mr. Ashish Guwalani, ACERC, Ajmer, India

TPC Chairs
Prof. Nitin Singh, Motilal Nehru National Institute of Technology Allahabad, India
Dr. Vishal Bhatnagar, AIACTR, Delhi, India

TPC Co-Chair
Dr. Sanjay Mathur, ACERC Ajmer, India

Publication Chairs
Dr. Deepak Kumar Singh, Sachdeva Institute of Technology, Mathura, India
Dr. Pragya Dwivedi, MNNIT Allahabad, India

Publication Co-Chair
Mr. Gaurav Phulwari, ACERC, Ajmer, India

Publicity Chairs
Dr. Anil Dubey, Government Engineering College, Ajmer, India
Dr. Deepak Kumar, Amity University, Noida, India
Dr. Nitin Rakesh, Amity University, Noida, India
Dr. Ravi Prasad Valluru, Narayana Engineering College Nellore, AP, India

Dr. Sushant Upadyaya, MNIT, Jaipur, India
Dr. Akshay Girdhar, GNDEC, Ludhiana, India

Publicity Co-Chair
Mr. Surendra Singh, ACERC, Ajmer, India

Tutorial Chairs
Prof. Lokesh Garg, Delhi College of Technology & Management, Haryana, India

Tutorial Co-Chair
Mr. Ankit Mutha, ACERC, Ajmer, India

Technical Program Committee

Prof. Ajay Gupta, Western Michigan University, USA
Prof. Babita Gupta, California State University, USA
Prof. Amit K.R. Chowdhury, University of California, USA
Prof. David M. Harvey, G.E.R.I., UK
Prof. Madjid Merabti, Liverpool John Moores University, UK
Dr. Nesimi Ertugrual, University of Adelaide, Australia
Prof. Ian L. Freeston, University of Sheffield, UK
Prof. Witold Kinsner, University of Manitova, Canada
Prof. Anup Kumar, M.I.N.D.S., University of Louisville, USA
Prof. Prabhat Kumar Mahanti, University of New Brunswick, Canada
Prof. Ashok De, Director, NIT Patna, India
Prof. Kuldip Singh, IIT Roorkee, India
Prof. A.K. Tiwari, IIT, BHU, Varanasi, India
Mr. Suryabhan, ACERC, Ajmer, India
Dr. Vivek Singh, BHU, India
Prof. Abdul Quaiyum Ansari, Jamia Millia Islamia, New Delhi, India
Prof. Aditya Trivedi, ABV-IIITM Gwalior, India
Prof. Ajay Kakkar, Thapar University, Patiala, India
Prof. Bharat Bhaskar, IIM Lucknow, India
Prof. Edward David Moreno, Federal University of Sergipe, Brazil
Prof. Evangelos Kranakis, Carleton University
Prof. Filipe Miguel Lopes Meneses, University of Minho, Portugal
Prof. Giovanni Manassero Junior, Universidade de São Paulo, Brazil
Prof. Gregorio Martinez, University of Murcia, Spain
Prof. Pabitra Mitra, Indian Institute of Technology Kharagpur, India
Prof. Joberto Martins, Salvador University-UNIFACS, Brazil
Prof. K. Mustafa, Jamia Millia Islamia, New Delhi, India
Prof. M.M. Sufyan Beg, Jamia Millia Islamia, New Delhi, India

Prof. Jitendra Agrawal, Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal, MP, India
Prof. Rajesh Baliram Ingle, PICT, University of Pune, India
Prof. Romulo Alexander Ellery de Alencar, University of Fortaliza, Brazil
Prof. Youssef Fakhri, Université Ibn Tofail, Faculté des Sciences, Brazil
Dr. Abanish Singh, Bioinformatics Scientist, USA
Dr. Abbas Cheddad, (UCMM), Umeå Universitet, Umeå, Sweden
Dr. Abraham T. Mathew, NIT, Calicut, Kerala, India
Dr. Adam Scmidit, Poznan University of Technology, Poland
Dr. Agostinho L.S. Castro, Federal University of Para, Brazil
Prof. Goo-Rak Kwon Chosun University, Republic of Korea
Dr. Alberto Yúfera, Instituto de Microelectrónica de Sevilla (IMSE), (CNM), Spain
Dr. Adam Scmidit, Poznan University of Technology, Poland
Prof. Nishant Doshi, S.V. National Institute of Technology, Surat, India
Prof. Gautam Sanyal, NIT Durgapur, India
Dr. Agostinho L.S. Castro, Federal University of Para, Brazil
Dr. Alok Chakrabarty, IIIT Bhubaneswar, India
Dr. Anastasios Tefas, Aristotle University of Thessaloniki
Dr. Anirban Sarkar, NIT-Durgapur, India
Dr. Anjali Sardana, IIIT Roorkee, Uttarakhand, India
Dr. Ariffin Abdul Mutalib, Universiti Utara Malaysia
Dr. Ashok Kumar Das, IIIT Hyderabad
Dr. Ashutosh Saxena, Infosys Technologies Ltd., India
Dr. Balasubramanian Raman, IIT Roorkee, India
Dr. Benahmed Khelifa, Liverpool John Moores University, UK
Dr. Björn Schuller, Technical University of Munich, Germany
Dr. Carole Bassil, Lebanese University, Lebanon
Dr. Chao Ma, Hong Kong Polytechnic University
Dr. Chi-Un Lei, University of Hong Kong
Dr. Ching-Hao Lai, Institute for Information Industry
Dr. Ching-Hao Mao, Institute for Information Industry, Taiwan
Dr. Chung-Hua Chu, National Taichung Institute of Technology, Taiwan
Dr. Chunye Gong, National University of Defense Technology
Dr. Cristina Olaverri Monreal, Instituto de Telecomunicacoes, Portugal
Dr. Chittaranjan Hota, BITS Hyderabad, India
Dr. D. Juan Carlos González Moreno, University of Vigo
Dr. Danda B. Rawat, Old Dominion University
Dr. Davide Ariu, University of Cagliari, Italy
Dr. Dimiter G. Velev, University of National and World Economy, Europe
Dr. D.S. Yadav, South Asian University, New Delhi
Dr. Darius M. Dziuda, Central Connecticut State University
Dr. Dimitrios Koukopoulos, University of Western Greece, Greece
Dr. Durga Prasad Mohapatra, NIT-Rourkela, India
Dr. Eric Renault, Institut Telecom, France
Dr. Felipe RudgeBarbosa, University of Campinas, Brasil

Dr. Fermín Galán Márquez, Telefónica I+D, Spain
Dr. Fernando Zacarias Flores, Autonomous University of Puebla
Dr. Fuu-Cheng Jiang, Tunghai University, Taiwan
Prof. Aniello Castiglione, University of Salerno, Italy
Dr. Geng Yang, NUPT, Nanjing, P.R. of China
Dr. Gadadhar Sahoo, BIT-Mesra, India
Prof. Ashokk Das, International Institute of Information Technology, Hyderabad, India
Dr. Gang Wang, Hefei University of Technology
Dr. Gerard Damm, Alcatel-Lucent
Prof. Liang Gu, Yale University, New Haven, CT, USA
Prof. K.K Pattanaik, ABV-Indian Institute of Information Technology and Management, Gwalior, India
Dr. Germano Lambert-Torres, Itajuba Federal University
Dr. Guang Jin, Intelligent Automation, Inc.
Dr. Hardi Hungar, Carl von Ossietzky University Oldenburg, Germany
Dr. Hongbo Zhou, Southern Illinois University Carbondale
Dr. Huei-Ru Tseng, Industrial Technology Research Institute, Taiwan
Dr. Hussein Attia, University of Waterloo, Canada
Prof. Hong-Jie Dai, Taipei Medical University, Taiwan
Prof. Edward David, UFS—Federal University of Sergipe, Brazil
Dr. Ivan Saraiva Silva, Federal University of Piauí, Brazil
Dr. Luigi Cerulo, University of Sannio, Italy
Dr. J. Emerson Raja, Engineering and Technology of Multimedia University, Malaysia
Dr. J. Satheesh Kumar, Bharathiar University, Coimbatore
Dr. Jacobijn Sandberg, University of Amsterdam
Dr. Jagannath V. Aghav, College of Engineering Pune, India
Dr. Jaume Mathieu, LIP6 UPMC, France
Dr. Jen-Jee Chen, National University of Tainan
Dr. Jitender Kumar Chhabra, NIT-Kurukshetra, India
Dr. John Karamitsos, Tokk Communications, Canada
Dr. Jose M. Alcaraz Calero, University of the West of Scotland, UK
Dr. K.K. Shukla, IT-BHU, India
Dr. K.R. Pardusani, Maulana Azad NIT, Bhopal, India
Dr. Kapil Kumar Gupta, Accenture
Dr. Kuan-Wei Lee, I-Shou University, Taiwan
Dr. Lalit Awasthi, NIT Hamirpur, India
Dr. Maninder Singh, Thapar University, Patiala, India
Dr. Mehul S. Raval, DA-IICT, Gujarat, India
Dr. Michael McGuire, University of Victoria, Canada
Dr. Mohamed Naouai, University Tunis El Manar and University of Strasbourg, Tunisia
Dr. Nasimuddin, Institute for Infocomm Research
Dr. Olga C. Santos, aDeNu Research Group, UNED, Spain

Dr. Pramod Kumar Singh, ABV-IIITM Gwalior, India
Dr. Prasanta K. Jana, IIT, Dhanbad, India
Dr. Preetam Ghosh, Virginia Commonwealth University, USA
Dr. Rabeb Mizouni, (KUSTAR), Abu Dhabi, UAE
Dr. Rahul Khanna, Intel Corporation, USA
Dr. Rajeev Srivastava, CSE, ITBHU, India
Dr. Rajesh Kumar, MNIT, Jaipur, India
Dr. Rajesh Bodade, Military College of Telecommunication, Mhow, India
Dr. Rajesh Kumar, MNIT, Jaipur, India
Dr. Ranjit Roy, SVNIT, Surat, Gujarat, India
Dr. Robert Koch, Bundeswehr University München, Germany
Dr. Ricardo J. Rodriguez, Nova Southeastern University, USA
Dr. Ruggero Donida Labati, Università degli Studi di Milano, Italy
Dr. Rustem Popa, University "Dunarea de Jos" in Galati, Romania
Dr. Shailesh Ramchandra Sathe, VNIT Nagpur, India
Dr. Sanjiv K. Bhatia, University of Missouri—St. Louis, USA
Dr. Sanjeev Gupta, DA-IICT, Gujarat, India
Dr. S. Selvakumar, National Institute of Technology, Tamil Nadu, India
Dr. Saurabh Chaudhury, NIT Silchar, Assam, India
Dr. Shijo M. Joseph, Kannur University, Kerala
Dr. Sim Hiew Moi, University Technology of Malaysia
Dr. Syed Mohammed Shamsul Islam, The University of Western Australia, Australia
Dr. Trapti Jain, IIT Mandi, India
Dr. Tilak Thakur, PED, Chandigarh, India
Dr. Vikram Goyal, IIIT Delhi, India
Dr. Vinaya Mahesh Sawant, D.J. Sanghvi College of Engineering, India
Dr. Vanitha Rani Rentapalli, VITS Andhra Pradesh, India
Dr. Victor Govindaswamy, Texas A&M University-Texarkana, USA
Dr. Victor Hinostroza, Universidad Autónoma de Ciudad Juárez
Dr. Vidyasagar Potdar, Curtin University of Technology, Australia
Dr. Vijaykumar Chakka, DAIICT, Gandhinagar, India
Dr. Yong Wang, School of IS & E, Central South University, China
Dr. Yu Yuan, Samsung Information Systems America—San Jose, CA
Eng. Angelos Lazaris, University of Southern California, USA
Mr. Hrvoje Belani, University of Zagreb, Croatia
Mr. Huan Song, SuperMicro Computer, Inc., San Jose, USA
Mr. K.K. Patnaik, IIITM, Gwalior, India
Dr. S.S. Sarangdevot, Vice Chancellor, JRN Rajasthan Vidyapeeth University, Udaipur
Dr. N.N. Jani, KSV University Gandhi Nagar, India
Dr. Ashok K. Patel, North Gujarat University, Patan, Gujarat, India
Dr. Awadhesh Gupta, IMS, Ghaziabad, India
Dr. Dilip Sharma, GLA University, Mathura, India
Dr. Li Jiyun, Donghua University, Shanghai, China

Dr. Lingfeng Wang, University of Toledo, USA
Dr. Valentina E. Balas, Aurel Vlaicu University of Arad, Romania
Dr. Vinay Rishiwal, MJP Rohilkhand University, Bareilly, India
Dr. Vishal Bhatnagar, Ambedkar Institute of Technology, New Delhi, India
Dr. Tarun Shrimali, Sun rise Group of Institutions, Udaipur, India
Dr. Atul Patel, CU Shah University, Vadhwan, Gujarat, India
Dr. P.V. Virparia, Sardar Patel University, VV Nagar, India
Dr. D.B. Choksi, Sardar Patel University, VV Nagar, India
Dr. Ashish N. Jani, KSV University Gandhi Nagar, India
Dr. Sanjay M. Shah, KSV University Gandhi Nagar, India
Dr. Vijay M. Chavda, KSV University Gandhi Nagar, India
Dr. B.S. Agarwal, KIT Kalol, India
Dr. Apurv Desai, South Gujarat University, Surat, India
Dr. Chitra Dhawale, Nagpur, India
Dr. Bikas Kumar, Pune, India
Dr. Nidhi Divecha, Gandhi Nagar, India
Dr. Jay Kumar Patel, Gandhi Nagar, India
Dr. Jatin Shah, Gandhi Nagar, India
Dr. Kamaljit I. Lakhtaria, Auro University, Surat, India
Dr. B.S. Deovra, B.N. College, Udaipur, India
Dr. Ashok Jain, Maharaja College of Engineering, Udaipur, India
Dr. Bharat Singh, JRN Rajasthan Vidyapeeth University, Udaipur, India
Dr. S.K. Sharma, Pacific University Udaipur, India
Dr. Akheela Khanum, Integral University Lucknow, India
Dr. R.S. Bajpai, Ram Swaroop Memorial University, Lucknow, India
Dr. Manish Shrimali, JRN Rajasthan Vidyapeeth University, Udaipur, India
Dr. Ravi Gulati, South Gujarat University, Surat, India
Dr. Atul Gosai, Saurashtra University, Rajkot, India
Dr. Digvijai sinh Rathore, BBA Open University Ahmadabad, India
Dr. Vishal Goar, Government Engineering College, Bikaner, India
Dr. Neeraj Bhargava, MDS University Ajmer, India
Dr. Ritu Bhargava, Government Women Engineering College, Ajmer, India
Dr. Rajender Singh Chhillar, MDU Rohtak, India
Dr. Dhaval R. Kathiriya, Saurashtra University, Rajkot, India
Dr. Vineet Sharma, KIET Ghaziabad, India
Dr. A.P. Shukla, KIET Ghaziabad, India
Dr. R.K. Manocha, Ghaziabad, India
Dr. Nandita Mishra, IMS Ghaziabad, India
Dr. Manisha Agarwal, IMS Ghaziabad
Dr. Deepika Garg, IGNOU New Delhi, India
Dr. Goutam Chakraborty, Iwate Prefectural University, Iwate Ken, Takizawa, Japan
Dr. Amit Manocha Maharaja Agrasen University, HP, India
Prof. Enrique Chirivella-Perez, University of the West of Scotland, UK
Prof. Pablo Salva Garcia, University of the West of Scotland, UK
Prof. Ricardo Marco Alaez, University of the West of Scotland, UK

Prof. Nitin Rakesh, Amity University, Noida, India
Prof. Mamta Mittal, G.B. Pant Government Engineering College, Delhi, India
Dr. Shashank Srivastava, MNNIT Allahabad, India
Prof. Lalit Goyal, JMI, Delhi, India
Dr. Sanjay Maurya, GLA University, Mathura, India
Prof. Alexandros Iosifidis, Tampere University of Technology, Finland
Prof. Shanthi Makka, JRE Engineering College, Greater Noida, India
Dr. Deepak Gupta, Amity University, Noida, India
Dr. Manu Vardhan, NIT Raipur, India
Dr. Sarsij Tripathi, NIT Raipur, India
Prof. Wg Edison, HeFei University of Technology, China
Dr. Atul Bansal, GLA University, Mathura, India
Dr. Alimul Haque, V.K.S. University, Bihar, India
Prof. Simhiew Moi, Universiti Teknologi Malaysia
Prof. Vinod Kumar, IIT Roorkee, India
Prof. Christos Bouras, University of Patras and RACTI, Greece
Prof. Devesh Jinwala, SVNIT Surat, India
Prof. Germano Lambert Torres, PS Solutions, Brazil
Prof. Byoungho Kim, Broadcom Corporation, USA
Prof. Aditya Khamparia, LPU, Punjab, India

Contents

Part III Intelligent Image Processing

About the Editors

Dr. Sanjiv K. Bhatia received his Ph.D. in Computer Science from the University of Nebraska, Lincoln in 1991. He presently works as Professor and Graduate Director (Computer Science) in the University of Missouri, St. Louis. His primary areas of research include image databases, digital image processing, and computer vision. He has published over 40 articles in these areas. He has also been consulted extensively by industry for commercial and military applications of computer vision. He is an expert in system programming and has worked on real-time and embedded applications. He serves on the organizing committee of a number of conferences and on the editorial board of international journals. He has taught a broad range of courses in computer science and was the recipient of Chancellor's Award for Excellence in Teaching in 2015. He is a senior member of ACM.

Dr. Krishn K. Mishra is currently works as a Visiting Faculty, Department of Mathematics and Computer Science, University of Missouri, St. Louis, USA. He is an alumnus of Motilal Nehru National Institute of Technology Allahabad, India, which is also his base working institute. His primary areas of research include evolutionary algorithms, optimization techniques, and design and analysis of algorithms. He has published more than 50 publications in international journals and proceedings of international conferences of repute. He has served as a program committee member of several conferences and also edited Scopus and SCI-indexed journals. He has 15 years of teaching and research experience during which he made all his efforts to bridge the gaps between teaching and research.

Dr. Shailesh Tiwari works as Professor in Computer Science and Engineering Department, ABES Engineering College, Ghaziabad, India. He is also administratively heading the department. He is an alumnus of Motilal Nehru National Institute of Technology Allahabad, India. He has more than 15 years of experience in teaching, research, and academic administration. His primary areas of research include software testing, implementation of optimization algorithms, and machine learning techniques in software engineering. He has also published more than 40 publications in international journals and in proceedings of international conferences of repute. He has served as a program committee member of several

conferences and edited Scopus and E-SCI-indexed journals. He has also organized several international conferences under the banner of IEEE and Springer. He is a Senior Member of IEEE, member of IEEE Computer Society, and Executive Committee member of IEEE Uttar Pradesh section. He is a member of reviewer and editorial board of several international journals and conferences.

Dr. Vivek Kumar Singh is Assistant Professor at Department of Computer Science, Banaras Hindu University, India. His major research interest lies in the area of text analytics. Currently, he is working on scientometrics, sentiment analysis, social network analysis, altmetrics, which are the broader research area of text analytics. He has developed and coordinated a text analytics laboratory, which works on various text analytics tasks. He is an alumnus of Allahabad University, Allahabad, India. He has published more than 30 publications in international journals and in proceedings of international conferences of repute. He has also served in South Asian University, Delhi, India as Assistant Professor for more than 4 years. He has also associated with several research projects such as Indo-Mexican Joint Research Project funded jointly by the Department of Science and Technology, Government of India, along with the National Council for Science and Technology (CONACYT) of the United Mexican States.

Part I
Advanced Software Engineering

Approach for an Opinion Wrapping System–Using Focused Web Crawler

Gaurav Vats, Vishal Bhatnagar, Rajat Sharma, Ishan Setiya and Arushi Jain

Abstract Most of the search engine depends on web crawler to go through a large number of Webpages. Web crawler (i.e. web spider or scutter or bot) is used to fetch content and URL from the Webpages. It also indexes them so that browser can easily and quickly fetch pages related to the searched word. Tons of data is produced every day, 90% of data has been created in the last 2 years. This data contains opinions and thoughts of the people in unstructured form. Opinion Scutter goes through the content and fetch reviews and comments so that they can be grilled and processed to find useful information. While shopping online or searching any game to buy we are largely dependent on the reviews provided by people. If we can keep track of such reviews and opinion, it will be easy to track a good product and increase efficiency and efficacy of the search engine. The proposed system is a generic crawler which fetches all the reviews from a given site.

Keywords Web crawler · Opinion mining · World Wide Web · Reviews · Webpage parser · Product monitoring

G. Vats (✉) · V. Bhatnagar · R. Sharma · I. Setiya · A. Jain
Ambedkar Institute of Advanced Communication Technologies and Research, New Delhi, India
e-mail: vats.gaurav101@gmail.com

V. Bhatnagar
e-mail: vishalbhatnagar@yahoo.com

R. Sharma
e-mail: rajvsrajat@gmail.com

I. Setiya
e-mail: setiya.ishan2781@gmail.com

A. Jain
e-mail: arushijain1391@gmail.com

S.K. Bhatia et al. (eds.), *Advances in Computer and Computational Sciences*, Advances in Intelligent Systems and Computing 554,
https://doi.org/10.1007/978-981-10-3773-3_1

1 Introduction

With the gigantic information present on Internet, we need an exceedingly proficient and systematic crawling system to parse through loads of websites. Exponential increase in the number of people that can access World Wide Web, there is significant growth of e-commerce. It has eased our life to a great extent, young generation prefers e-malls rather than traditional markets because of more options on their tip. With more options and offers it is hard to choose the best. The only thing on which we rely are the opinions of the people. E-commerce websites competes other on the basis of their reviews and ratting given by users. Those reviews are just a bunch of sentences which do not have any meaning as a cluster. In order to fetch them and further process them we have proposed an 'Opinion Parsing System'. The proposed system goes through the website and fetches all the reviews (opinion) and URLs. We can also do the analysis of the opinions collected.

2 Literature Review

Immense research is being done in the field of Opinion Mining and Focused Crawling but there lie a difference in every approach. Generally, these approaches are performed in a two-step execution model.

Song et al. [1] used the approach of Focused Crawling for collecting relevant pages from the web for a given topic and extracts the pages that contain opinions. Sentiment analysis is then performed on these pages to predict the opinion's polarity. It also finds out the latest webpages related to a given topic by making use of Agents.

Shkapenyuk et al. [2] provided a distributed crawler system which is highly flexible and achieves high performance by crawling quite a lot of pages within a second. The work proposed by them is divided into two mechanisms-: Crawler Application and Crawler System. The Crawling System helps in downloading a large number of web pages within some seconds and Crawling Application uses techniques such as Breadth First Crawl. Crawl Manger is the main constituent of the Crawler Application.

Dhawan et al. [3] proposed a system that evaluates different domains of information on twitter such as sports, etc., and expand these domains into further levels based on the score they achieved through every tweet in the particular domain.

Kang et al. [4] proposed a model known as wrapper model for collecting data from e-commerce websites. He further divided this data into two different aspects, the one that contain reviews and the later one that do not contain reviews. It stores the collected data in a table. Kang made use of four analysis processes to carry out its implementation.

The approach which Mfenyana et al. [5] used was a relatively different one. They proposed the Facebook crawler using Iterative Incremental development approach. This paper aims to determine the opinions trending on Facebook about a

specific topic using the technique called index searching. Frequency Analysis Module used in this approach is a remarkable research as it provides information about the repetitive behavior of a particular topic.

Nowadays, a large number of opinions or reviews can be found on the e-commerce websites because of the rapid growth and development in the ecommerce. Yahui Xi [6] aim to focus mainly on review mig and product features exaction. They used the technique called Double Propagation to improve precision and recall.

Pappas et al. [7] proposed a unique system which would be able to discover the trending topics on the web pages related to any blog, discussions or news which must include opinionated texts or reviews of the users. The methodology used by them was focused crawling as they likely to bind to a particular topic such as news. Web segmentation is done on these pages to spot regions generated by user's opinions. A confidence score is further calculated which computes the percentage on which the given topic is related to the page and finally performs sentiment polarity on the regions of user generated text.

3 Proposed Methodology

All the existing crawler have certain problems we have introduced this concept to increase the efficiency of the opinion parser, using crawler. This is a generic architecture that can be implemented by any web bot which fetches any opinions, reviews, comments, blogs, etc. (as can be seen in Fig. 1 and Table 1).

The Seed URL is given by the user, it is the starting point from where the crawler will start working and is the first page to be parsed. This page is parsed by URL RETRIEVER, it searches out all the URLs present in that page and every page

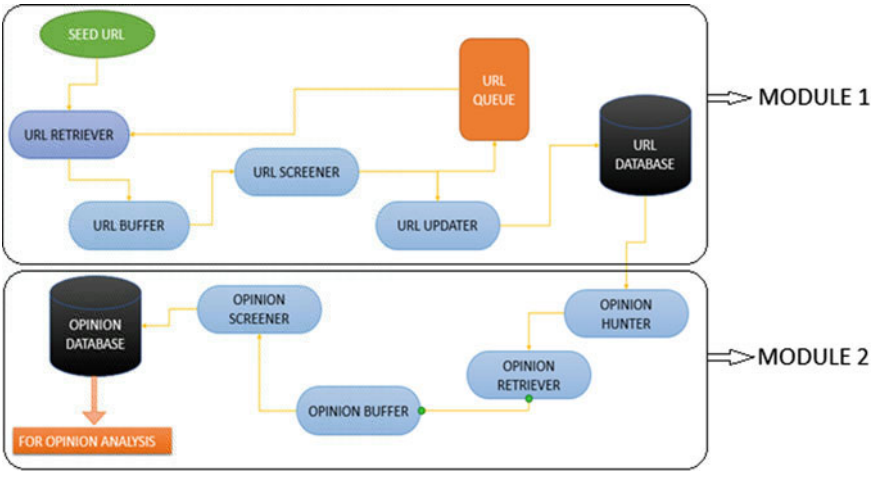

Fig. 1 Opinion crawler architecture

Table 1 Architecture description

1.	SEED URL	It is the main URL given by the user
2.	URL RETRIEVER	It fetches all the URL on the Page and stores it in buffer
3.	URL BUFFER	It holds all the unparsed URLs, which
4.	URL SCREENER	It filters the URLs which might not contain the reviews or comments
5.	URL UPDATER	Updates the URL to database and processing queue
6.	URL DATABASE	Stores URLs, along with URL id, timestamp of last visited
7.	OPINION HUNTER	It tracks the opinion or reviews on a page
8.	OPINION RETRIEVER	It fetches the opinion from the pages which are embedded in different elements of HTML/DHTML
9.	OPINION SCREENER	It filters all the duplicate and false reviews
10.	OPINION DATABASE	It stores all the opinion fetched by scutter, along with the product id, user id, last updated and ratting

given to it as parameter. All the URLs fetched by the retriever go into the URL BUFFER, which stores all the pending URLs for screening by the URL SCREENER. This block filters out all the foreign links one by one from the buffer. Rest of the links which passed the screening is updated to URL DATABASE by URL UPDATER. The updater only updates links which may contain reviews or comments. These URLs are parsed by OPINION HUNTER which hunts down the location of opinions in the page. Opinions from these location are fetched by the OPINION RETRIEVER. The comments fetched are further filtered out to reduce the redundancy of the opinion by the same person. The opinions which remain are added to the OPINION DATABASE. These opinions/comments/reviews can be further used for analysis (as shown in Fig. 1).

4 Proposed Algorithms

We authors used different algorithms for different blocks and implemented them in JAVA.

4.1 URL Retriever

This block retrieves all the URLs present on the seed URL and after that every page's URL provided to and forward all the URLs extracted to the URL Buffer. This block get URLs form the source code provided to browser. It uses predefined API 'Jsoup'.

```
URL_Retriever(String URL){
     1 Search for all URL on page
          Using JSOUP Srapper API
          Elements questions = doc.select("a[href]");
     2 Check if Current URL is already in queue
     3 Send URL to Buffer
     4 Call URL_Screener();
     }
```

4.2 URL Screener

As all the URLs retrieved by the URL Retriever will not be pages, which contains reviews or even links to different sites such as Facebook or Gmail or different advisements page/links. These links should be filtered out so that we do not have to parse them. This will increase the efficiency of crawling system as when we check for any update we will directly go to the pages which contain products. This is implemented using sting matching in URL. As URLs which do not have the website name will be filtered out, for e.g. every link of eBay will contain String “eBay’’ in it.

```
URL_Screener (String URL){
     1 Filter the seedURL to get filteredURL
          String filteredURL = URLfilter (seedURL)
     2 for (Element Link: questions{
     3 //check if URL belongs to website
          if(Link.attr("href").contains(filteredURL))
     4     call URL_Updater(Link);
     }  }
```

4.3 URL Updater

This block updates the filtered URLs to the URL Queue. The updater put the parsed URL to the database putting timestamp on it and the URLs fetched from that page to the URL Queue.

```
URL_Updator(String URL){
     1 Parse all content of current web page
          Document doc =  Jdoup.connect(URL).get();
     2 Check weather webpage is  relevent
          if(do.text().contains(Reviews))
     3 Update link to URL database
     4 call URL_retriever();
     }
```

4.4 Opinion Hunter

This block then get the URLs one by one and hunt for the part which contains the opinion or reviews of the people. As there is no standard rule for making webpages,

Table 2 Example of knowledge table

Domain name	Class name
amazon.in	reviewText
Flipkart.com	rightCol
Tripadvisor.in	Entry

languages like JavaScript and Ajax makes it hard to track the comment or review on different pages. We observed that most website use CSS and the review section is contained and formatted using div tag with specific class, so to make a generalized System we used Knowledge Table (KT) in algorithm. KT is a table, which is formed with pair of domain name and class name of div containing review particular websites (see Table 2).

KT is implemented using hashing so searching time is O(1) which make it Generalized algorithm with better efficiency

```
Opinion_Hunter(){
    1 Providing A Knowledge_Table(KT) which contains
        Knowledge about ClassName of Div of Particular
        websites
    2 Find required ClassName Usong above KT
        String classNAme = compare(KT.URL, FilteredURL);
    3 Extract Reviews in String
        Elements divs = doc.select("div.className");
        for(Elements d : divs){
    4 Store d.text() in Opinion Buffer
    5 call Opinion Screener
    6 Store in Opinion Database
    }          }
```

4.5 Opinion Screener

This block screens out reviews on the basis of redundancy, only one review from one customer will be entertained. There are people who post more than one reviews to increase the overall rating or to degrade the rival's product rating. Latest comment will be considered by each user. This will prevent our analysis from opinion warfare to some extent and the review will be genuine.

```
Opinion_Screener(){
    1 Check whether the customer have earlier commented
        on that product.
    2 If record exists for that particular customer
        product pair then update the record.
    3 Else enter a new record to Opinion Database.
    }
```

Our proposed model and algorithms will decrease the time taken by the wrapping crawler. As we can deduce from the algorithm:

For a generic crawler time complexity can be written as

$O(u^a)$; *where u = total no. URLs & a = total anchor tags*

For our crawler time complexity will be

$O(u^{a'} - K) = O(u^{a'})$; *where u = total no. URLs & a' = a-total irrelevant tags; K = Keywords*