

Lecture Notes on Data Engineering and Communications Technologies

Volume 16

Series editor

Fatos Xhafa, Technical University of Catalonia, Barcelona, Spain
e-mail: fatos@cs.upc.edu

The aim of the book series is to present cutting edge engineering approaches to data technologies and communications. It publishes latest advances on the engineering task of building and deploying distributed, scalable and reliable data infrastructures and communication systems.

The series has a prominent applied focus on data technologies and communications with aim to promote the bridging from fundamental research on data science and networking to data engineering and communications that lead to industry products, business knowledge and standardisation.

More information about this series at <http://www.springer.com/series/15362>

Durgesh Kumar Mishra · Xin-She Yang
Aynur Unal
Editors

Data Science and Big Data Analytics

ACM-WIR 2018

Editors

Durgesh Kumar Mishra
Department of Computer Science
and Engineering
Sri Aurobindo Institute of Technology
Indore, Madhya Pradesh
India

Aynur Unal
Department of Mechanical Engineering
Indian Institute of Technology Guwahati
Guwahati, Assam
India

Xin-She Yang
School of Science and Technology
Middlesex University
London
UK

ISSN 2367-4512 ISSN 2367-4520 (electronic)
Lecture Notes on Data Engineering and Communications Technologies
ISBN 978-981-10-7640-4 ISBN 978-981-10-7641-1 (eBook)
<https://doi.org/10.1007/978-981-10-7641-1>

Library of Congress Control Number: 2018941227

© Springer Nature Singapore Pte Ltd. 2019

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.

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.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

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

Preface

The pervasive nature of digital technologies as witnessed in industry, services, and everyday life has given rise to an emergent, data-focused economy stemming from many aspects of human individual and commercial activity. The richness and vastness of these data are creating unprecedented research opportunities in some fields including urban studies, geography, economics, finance, and social science, as well as physics, biology and genetics, public health, and many others. In addition to big data-inspired research, businesses have seized on big data technologies to support and propel growing business intelligence needs. As businesses build out big data hardware and software infrastructure, it becomes increasingly important to anticipate technical and practical challenges and to identify best practices learned through experience. Big data analytics employ software tools from advanced analytics disciplines such as data mining, predictive analytics, and machine learning. At the same time, the processing and analysis of big data present methodological and technological challenges. The goal of this lecture series and symposium is to present both novel solutions to challenging technical issues and compelling big data use cases. This special issue contains many papers that provide deep research results to report the advances in big data analytics, infrastructure, and applications. The goal of this special issue is to crystallize the emerging big data technologies and efforts to focus on the most promising solutions in the industry. The papers provide clear proof that big data technologies are playing a more and more important and critical role in supporting various applications in the industry. It is also believed that the papers will further research openings in new best practices and directions in this emerging research discipline.

Indore, India
London, UK
Guwahati, India

Durgesh Kumar Mishra
Xin-She Yang
Aynur Unal

Contents

A Study of the Correlation Between Internet Addiction and Aggressive Behaviour Among the Namibian University Students	1
Poonam Dhaka and Cynthia Naris	
Genetic Algorithm Approach for Optimization of Biomass Estimation at LiDAR	11
Sonika and Aditi Jain	
E-alive: An Integrated Platform Based on Machine Learning Techniques to Aware and Educate Common People with the Current Statistics of Maternal and Child Health Care	29
Garima Malik, Sonakshi Vij, Devendra Tayal and Amita Jain	
An Effective TCP's Congestion Control Approach for Cross-Layer Design in MANET	43
Pratik Gite, Saurav Singh Dhakad and Aditya Dubey	
Baron-Cohen Model Based Personality Classification Using Ensemble Learning	57
Ashima Sood and Rekha Bhatia	
Investigation of MANET Routing Protocols via Quantitative Metrics	67
Sunil Kumar Jangir and Naveen Hemrajani	
Review on Internet Traffic Sharing Using Markov Chain Model in Computer Network	81
Sarla More and Diwakar Shukla	
Anomaly Detection Using Dynamic Sliding Window in Wireless Body Area Networks	99
G. S. Smrithy, Ramadoss Balakrishnan and Nikita Sivakumar	

Effective Healthcare Services by IoT-Based Model of Voluntary Doctors	109
Bharat B. Prajapati, Satyen M. Parikh and Jignesh M. Patel	
Emotional State Recognition with EEG Signals Using Subject Independent Approach	117
Pallavi Pandey and K. R. Seeja	
Development of Early Prediction Model for Epileptic Seizures	125
Anjum Shaikh and Mukta Dhopeshwarkar	
Research Issue in Data Anonymization in Electronic Health Service: A Survey	139
Amanuel Gebrehiwot and Ambika V. Pawar	
Prediction of Cervical Cancer Based on the Life Style, Habits, and Diseases Using Regression Analysis Framework	149
K. Vinoth Raja and M. Thangamani Murugesan	
Novel Outlier Detection by Integration of Clustering and Classification	169
Sarita Tripathy and Laxman Sahoo	
A Review of Wireless Charging Nodes in Wireless Sensor Networks	177
Shiva Prakash and Vikas Saroj	
Leeway of Lean Concept to Optimize Big Data in Manufacturing Industry: An Exploratory Review	189
Hardik Majiwala, Dilay Parmar and Pankaj Gandhi	
Monitoring Public Participation in Multilateral Initiatives Using Social Media Intelligence	201
Ulanat Mini, Vijay Nair and K. Poullose Jacob	
An Efficient Context-Aware Music Recommendation Based on Emotion and Time Context	215
C. Selvi and E. Sivasankar	
Implementation of Improved Energy-Efficient FIR Filter Using Reversible Logic	229
Lavisha Sahu, Umesh Kumar and Lajwanti Singh	
A Study on Benefits of Big Data for Healthcare Sector of India	239
Komal Sindhi, Dilay Parmar and Pankaj Gandhi	

Handling Uncertainty in Linguistics Using Probability Theory	247
Annapurna P. Patil, Aditya Barsainya, Sairam Anusha, Devadas Keerthana and Manoj J. Shet	
Review of Quality of Service Based Techniques in Cloud Computing	255
Geeta and Shiva Prakash	
Skyline Computation for Big Data	267
R. D. Kulkarni and B. F. Momin	
Human Face Detection Enabled Smart Stick for Visually Impaired People	277
Shivam Anand, Amit Kumar, Meenakshi Tripathi and Manoj Singh Gaur	
Web-Based Service Recommendation System by Considering User Requirements	291
Neha Malviya and Sarika Jain	
Unsupervised Machine Learning for Clustering the Infected Leaves Based on the Leaf-Colors	303
K. Ashok Kumar, B. Muthu Kumar, A. Veeramuthu and V. S. Mynavathi	
Real-Time Big Data Analysis Architecture and Application	313
Nandani Sharma and Manisha Agarwal	
Missing Value Imputation in Medical Records for Remote Health Care	321
Sayan Das and Jaya Sil	
Recommendation Framework for Diet and Exercise Based on Clinical Data: A Systematic Review	333
Vaishali S. Vairale and Samiksha Shukla	
Security Assessment of SAODV Protocols in Mobile Ad hoc Networks	347
Megha Soni and Brijendra Kumar Joshi	
Secure Sum Computation Using Homomorphic Encryption	357
Rashid Sheikh and Durgesh Kumar Mishra	
Automated Workload Management Using Machine Learning	365
K. Deivanai, V. Vijayakumar and Priyanka	
Multi-user Detection in Wireless Networks Using Decision Feedback Signal Cancellation	379
Monika Sharma and Balwant Prajapat	

ANN-Based Predictive State Modeling of Finite State Machines	389
Nishat Anjum and Balwant Prajapat	
Deep Dive Exploration of Mixed Reality in the World of Big Data	401
Prajat Mishra	
Author Index	405

About the Editors

Dr. Durgesh Kumar Mishra is a professor (CSE) and director of the Microsoft Innovation Centre at Sri Aurobindo Institute of Technology, Indore, India, and visiting faculty at IIT Indore. He has 24 years of teaching and 12 years of research experience. He has published more than 90 papers in refereed international/national journals and conferences including IEEE, ACM and organized many conferences as General Chair and Editor. He is a senior member of IEEE, CSI, ACM, chairman of IEEE MP subsection, IEEE Computer Society Bombay Chapter. At present, he is the chairman of CSI Division IV Communication at the National Level and ACM Chapter Rajasthan and MP State.

Prof. Xin-She Yang is an associate professor of Simulation and Modelling at Middlesex University, London. His main interests are applied mathematics, algorithm development, computational intelligence, engineering optimization, mathematical modeling, optimization, and swarm intelligence. His research projects have been supported by the National Measurement Office, BIS, Southwest Development Agency (UK), Euro Met, EPSRC, NPL, and the National Science Foundation of China. He is EEE CIS Task Force Chair of the BIKM, Technical Committee of Computational Finance and Economics of IEEE Computational Intelligence Society; Advisor to the *International Journal of Bio-Inspired Computation*; Editorial Board Member of *Elsevier's Journal of Computational Science*; and Editor-in-Chief of the *International Journal of Mathematical Modelling and Numerical Optimisation*.

Dr. Aynur Unal is a strategic adviser and visiting full professor at the IIT Guwahati, India. She has created a product-focused engineering program using the cloud-based infrastructure. Her main interests include ecologically and socially responsible engineering, zero waste initiative, and sustainable green engineering. Her research focuses on both rural and urban sustainable development, renewable

energy, solar towers and pumps. She has taught at Stanford University and worked in Silicon Valley to develop products for data mining from big data (Triada's Athena I & II), collaborative design and manufacturing, secure and private communication, and collaboration software platforms (Amteus, listed in LSE AIM).

Keynote Speakers

Analysis and Evaluation of Big Data Video Quality in Wireless Networks

Talks Abstract

With the development of the Internet technologies and online multimedia applications of the Internet, video becomes the main source of online-generated data. To measure the quality of received video, computer simulation is the research tool of choice for a majority of the wired and wireless network research community. These days, most of the research works on the network adopt computer simulation to verify novel ideas. The presentation introduces a framework and a tool-set for evaluating the quality of video transmitted over a simulated wireless network. Besides measuring the quality of service (QoS) parameters of the underlying network, such as loss rates, delays, and jitter, frameworks also support a video quality evaluation of the received video. On wired as well as wireless networks, the medium access control (MAC) plays an important role in the performance of video transmission. MAC handles allocating the resources to the different types of applications or wireless stations. Many types of researches are conducted for video quality measurement on network and application layer. The framework can be used for research and evaluating new techniques for MAC layer optimizations. This talk will present an overview of the framework and a tool-set for evaluating the quality of video transmitted over a simulated wireless network. Finally, the future research directions will be discussed.

Keywords Big data video, Networks and communications, QoS, Multimedia technologies



Dr. Dharam Singh Jat ACM Distinguished Speaker and Professor, Department of Computer Science, Namibia University of Science and Technology, Windhoek, Namibia.

Dharm Singh Jat received his degrees Master of Engineering and Ph.D. in Computer Science and Engineering from prestigious universities in India. He is a professor in the Department of Computer Science at Namibia University of Science and Technology. From 1990 to 2014, he was with College of Technology and Engineering, MPUAT, Udaipur, India, and has more than 27 years of academic experience. He was instrumental in the setting up of MPUAT Intranet. He has given several guest lectures/invited talks at various prestigious conferences such as 45th Annual National Convention of Computer Society of India (2010), 26th Indian Engineering Congress at Bangalore (2011), and various ACM conferences as ACM speaker. More than 8 Ph.D. and 30 Master students have been guided by him who are occupying leading positions including professor in academia and industry. His interests span the areas of multimedia communications, wireless technologies, mobile communication systems, roof computing, and video transmission over wired–wireless networks, software-defined networks, network security, Internet of things, and ICT Applications.

Publishing Ethics and Author Services

Talks Abstract

The importance of research publishing can be defined by a simple quote of Gerard Piel, which says “Without publication, science is dead.” The first scientific journal was published in 1665, and we have traveled 350 years since then. In the last 20 years, science and reporting of science have undergone revolutionary changes. Computerization and the Internet have changed the traditional ways of reading and writing. Hence, it is very important for scientists and students of the sciences in all

disciplines to understand the complete process of writing and publishing of a scientific paper in good journals. There is also a downside of digital publishing. The principal challenge for publishers is to handle ethical issues, and it is of utmost importance for the authors to understand the ethical practices involved in the process.

Keyword Author services for the publishing work, Computerization and science



Mr. Aninda Bose Senior Publishing Editor with Springer India Pvt. Ltd.

Aninda Bose is presently working as a senior publishing editor with Springer India Pvt. Ltd. He is part of the Global Acquisition Team at Springer and responsible for acquisition of scientific content across the globe. The disciplines he is responsible for are chemistry, material sciences, earth sciences, and computational intelligence. He has more than 23 years of industrial experience in marketing and different fields of publishing. He has completed Master's in organic chemistry from Delhi University and Master's in marketing research from Symbiosis Institute of Management Studies, Pune. He has published books for secondary level in chemistry and a member of American Chemical Society, USA. He has delivered more than 50 talks on Scientific Writing and Publishing Ethics in reputed universities, international conferences, and author workshops.

Applying Machine Learning Techniques for Big Data Analytics

Talks Abstract

Big data is gaining tremendous importance in today's digital world. Big data is a collection of data sets that are so voluminous and complex that traditional data processing application software seems inadequate to deal with them. This data is diverse in nature and grows exponentially. It includes both organized and

unorganized formats. Some of the data collected from the transactions like credit/debit cards, e-commerce database, social networking sites, patient data from hospitals, and student's data from MOOCs/colleges could be stated as examples of big data. Handling this data type of data and more importantly analyzing this data to predict the near future is a tedious task for any analyst. Predictive analytics is a technique that uses both new and historical data to forecast future activity, behavior, and trends. It involves applying statistical analysis techniques, analytical queries, and predictive algorithms to the data sets to create models that show a numerical value, or score on the likelihood of a particular event happening. Predictive analytics deals with extracting the patterns from the data available so that it would be presented in a useful way to the concerned people, be it in businesses, health care, or any other field. As the domain changes, the properties of the data set also changes, and hence, different predictive models seem to be suitable for different domains.

In our present research, we have applied many machine learning techniques and tools. The talk would focus on the experimental results obtained in some of the domains stated as above.



Dr. Annapurna P. Patil Professor, CSE, M.S. Ramaiah Institute of Technology, Bangalore, India.

Dr. Annapurna P. Patil is working as a professor in the Department of Computer Science and Engineering at M.S. Ramaiah Institute of Technology, Bangalore, India. She has an experience of 23 years in academics. She is a senior IEEE member, LMCSI, LMISTE, member ACM, faculty advisor of IEEE, WiE, MSRIT. She has published several papers in various national and international conferences and in reputed journals and chaired sessions. Her teaching interest spans areas of mobile ad hoc networks, protocol engineering, artificial intelligence, and predictive models for data analytics, advanced algorithms, software engineering, distributed computing, bio-inspired computing, IoT, and cloud.

She has been involved in collaborative works with industries like CISCO, IBM, HPE, TransNeuron and Nihon Communications Ltd, Bangalore, in the areas of wireless networks, cognitive computing, IoT, and cloud.

Outlier Detection in Big Data

Talks Abstract

According to Hawkins (Hawkins 1980): “An outlier is an observation which deviates so much from the other observations as to arouse suspicions that it was generated by a different mechanism.” Outlier detection extends traditional database monitoring with increased intelligence that helps security analysts understand risk based on relative change in behavior. Finding outliers may generate high value if they are found, the value in terms of cost savings, improved efficiency, compute time savings, fraud reduction, and failure prevention. According to IBM statistics, “Every day, we create 2.5 quintillion bytes of data—so much that 90% of the data in the world today has been created in the last two years alone. This data comes from everywhere: sensors used to gather climate information, posts to social media sites, digital pictures and videos, purchase transaction records, and cell phone GPS signals to name a few. This data is big data.” These large-scale heterogeneous types of data appear problematic for traditional outlier detection methods to process. In this talk, we will discuss some of the most relevant outlier detection techniques for big data, research issues, and its applications.



Dr. Ramadoss Balakrishnan Professor, National Institute of Technology, Tiruchirappalli.

Dr. (Mr.) Ramadoss Balakrishnan received M.Tech. degree in Computer Science and Engineering from IIT Delhi in 1995 and earned Ph.D. degree in Applied Mathematics from IIT Bombay in 1983. Currently, he is working as a professor of Computer Applications at National Institute of Technology, Tiruchirappalli. His

research interests include software testing methodologies, security and privacy in big data and cloud, data mining, predictive analytics in big data, and multimedia mining.

Under his guidance, eight candidates have successfully completed Ph.D. and five candidates are pursuing Ph.D. He has more than 35 refereed international journal publications and more than 40 papers presented in reputed international conferences.

He is a recipient of Best Teacher Award at National Institute of Technology, Tiruchirappalli, India, during 2006–2007. He is a member of IEEE; Life Member (LM) of ISTE, New Delhi; Life Member (LM), Computer Society of India. For more details, visit <https://www.nitt.edu/home/academics/departments/faculty/Brama.pdf>.

Identification and Prevention of Internet Addiction in Youth

Talks Abstract

With the advancement of the Internet technology and available resources, the accessibility and attractiveness of youth toward using Internet have been enhanced. It has an impact on individual's cognitive learning, physical and mental development for adolescents. Addiction to the Internet can negatively impact family relationships and adolescents' behavior. Early stage identification of Internet addiction and its risk factor is therefore a clinical significance for the prevention of the Internet addiction in youth. This presentation explores the impact of level of the Internet addiction and aggression as well as the prevalence of the different forms of aggression in youth. Addiction to the Internet causes behavioral and emotional problems. Teenagers and youth feel restless if they are restricted to use the Internet. For reducing behavioral problems, counseling seems to be a significant method.

Keywords Internet addiction, Psychological well-being, Social networks, Internet addiction prevention



Poonam Dhaka Department of Human Sciences, University of Namibia, Windhoek, Namibia.

Dr. Poonam Dhaka received her Ph.D. degree in Psychology from MLSU, India, and Master of Science in Home Science from College of Home Science, Udaipur, India. She has more than 20 years of academic and research experience, and presently she is a senior lecturer in clinical psychology with the Department of Human Sciences, University of Namibia. Her interests span the areas of mind power techniques, cyberpsychology and behavior, health psychology, mental health disorders, psychological assessment, gender identity, cognitive psychology, parent-child relationships. She has developed clinical psychological tools and mind power techniques for teenagers to increase intelligence, creativity, and learning. She has authored and co-authored over 37 research publications in peer-reviewed reputed journals, chapters, and conference proceedings. She is a member of Psychology Association of Namibia, Computer Society of India and served as a member of various committees in the national and international events.

Android: Security Issues

Talks Abstract

Android provides built-in security at the kernel level and framework level. Android Linux kernel is responsible for provisioning application sandboxing and enforcement of some Linux permissions. Sandboxing provides isolation where each application runs within its own user ID and virtual machine. At framework level, system-built and user-defined permissions make the access to software and hardware resources restricted. Android also provides app signing where installation of only signed apps is permissible. In spite of built-in security features, malware developers have succeeded in implementing various attacks on Android. One of the reasons is mistake made by the end user who normally grants permissions without

any judgement. The other reasons are Android's architecture which permits app repackaging, dynamic code loading, reflection, use of native code, etc. Some intelligent malware writers utilize covert channels also to perform attacks. These attacks include stealing of private information, resource drainage, frauds, premium-rate SMSs.

Our work is focused on meliorating detection of malware that uses reflection as the means of hiding leaks. Reflection is a programming language feature that permits analysis and transformation of the behavior of classes used in programs in general and in apps in particular at the runtime. Reflection facilitates various features such as dynamic class loading, method invocation, and attribute usage at runtime. Unfortunately, malware authors leverage reflection to subvert the malware detection by static analyzers as reflection can hinder taint analysis used by static analyzers for the analysis of sensitive leaks. Even the latest and probably the best performing static analyzers are not able to detect information leaks in the malware via reflection. We proposed a system that combines dynamic analysis with code instrumentation for a more precise detection of leaks in malicious apps via reflection with code obfuscation. The evaluation of approach shows substantial improvement in detection of sensitive leaks via reflection.



Dr. Meenakshi Tripathi Associate Professor, Department of Computer Science, MNIT, Jaipur, India.

Dr. Meenakshi Tripathi has more than 10+ years of experience. She has completed M.Tech. degree in 2005 at Computer Science and Engineering, Banasthali Vidyapith, Banasthali, and Ph.D. degree in 2014 at Computer Science and Engineering, MNIT, Jaipur. Her areas of expertise are wireless sensor networks, information and network security, software-defined networks, and Internet of things. She has supervised 5 Ph. D. scholars and more than 15 master's thesis. She is handling Department of Science and Technology, sponsored R&D projects. Her 20+ publications are in peer-reviewed international journal and conferences. She is the principal investigator in DST projects having around 60 lakhs of funding.

Tweaking Big Data Analysis for Malicious Intent

Talks Abstract

Governments of various nations and major industries working in the world have started relying on big data analysis for future predictions of their workflow and making critical decisions based on the results. With the indefinite scope of big data in today's scenario, assessing the risks pertaining to data breach, data tampering, and all related security factors is at a very minimal level. Understanding all the malicious activities and intents from the tweaking of big data analysis by individuals, organizations, or governments that could lead to catastrophic outcomes should be treated with utmost priority.



Mr. Prajal Mishra President of Graduate Student Senate at University of Texas at Arlington, Arlington, TX, USA.

Prajal is a graduate student at University of Texas, Arlington, pursuing his Master of Science in Computer Science, specializing in the field of databases and software engineering. He also focuses on secure programming paradigms and will be graduating in Spring 2018. He is working as the President of Graduate Student Senate at University of Texas at Arlington, which is the official representative body of all graduate students on campus. During his undergraduate, he was a Microsoft Student Partner for 2 years and co-organized “Hour of Code, Indore” in the year 2014 and 2015 with a team of 60+ volunteers who taught and motivated 5000+ school students from 25+ schools about coding.

Big Data and Web Mining

Talks Abstract

The era of growing Web users on the Internet, i.e., the world of the World Wide Web, has led to the generation of massive data with such a high speed and variety that it has given a scope to analyze the behavior of Web user. Almost all the streams existing today have become almost dependent on Web technologies, and thus, we have the opportunity to analyze the user data which can help us in improvising services and the products offered, utilizing the likes and dislikes of the user to ensure a better user experience and by providing better quality products. Predicting the trends of user can be helpful in deriving his future patterns. The user's past data can be helpful in drawing out this pattern. This is one of the paradigms where software products can be harnessed to provide more pleasing experience to the users.

Big data tools and technologies help to mine patterns, association, prediction for improving quality of service, applying personalization, i.e., recommendation of product or promotional offers, placing advertise, to handle business-specific issues like customer attraction, retention, cross sales, and departure. Web mining uses Web content, hyperlinks, server logs, and other Web data to retrieve useful information and draw the patterns. Big data challenges first deals with the massive volume of data which includes Web logs, content, their social media data like Facebook, Twitter, images, video. Fortunately, advance data management like cloud computing and virtualization provides effective capture, storage, and manipulation of the large volume of data. Secondly, it helps in dealing with a variety which includes structured data like the product information, unstructured data like the tweet by the user, Web page content, and the semi-structured data. Big data tools and techniques have the potential to convert, store, access, analyze, and visualize a variety of data. The continuous flow of massive new data, i.e., the data velocity, happens to be a bigger challenge today. So, the volume and variety of data that are collected and stored have changed. Real-time data is necessary to retrieve, analyze, compare, and decide the user trends and shall totally depend on the new findings offered by the big data technologies.



Dr. Bhawna Nigam Assistant Professor, IET, Devi Ahilya University, Indore, India.

Dr. Bhawna Nigam received her B.E. and M.E. degrees with honors in Computer Engineering Institute of Engineering and Technology (IET), Devi Ahilya University, Indore, India, in 2003 and 2008, respectively. In 2017, she obtained her Ph.D. degree in Computing Engineering. She is currently with Institute of Engineering and Technology (IET), Devi Ahilya University, Indore, India, as an assistant professor in Information Technology Department since 2007. Her current research interests include data mining, Web mining, big data analytics, machine learning. She has published 20+ papers.

Big Data and Data Science Trends: Challenges and Opportunities

Talks Abstract

Due to rapid growth in data, our society is facing a paradigm shift in Information and Communication Technology (ICT). These trends are going to change the industry and business process and impact the society. The basis of this paradigm shift is data science which gives meaning to the big data. Data science framework consists of three basic elements: people, data, and technology. To deal with data science framework, data scientist should possess few important characteristics such as data-driven approach, computational intelligence, and latest technological knowledge for better analysis, visualization, and preservation of huge amount of data. Here, we will discuss various digital challenges and opportunities ahead of us in big data and data science.

Keywords Data science, Big data, Trends, Challenges, Opportunity, ICT



Dr. Samiksha Shukla Assistant Professor, Department of Computer Science and Engineering, Christ University, Bengaluru, Karnataka, India.

Dr. Samiksha Shukla received her M.Tech. degree in Computer Science and Engineering from RGTU, Bhopal, and Ph.D. degree in Computer Science Engineering from Christ University, Bengaluru, India. She is currently working as an assistant professor in Department of Computer Science and Engineering, Christ University. She has published 35 papers in peer-reviewed conferences and journals in the area of data and computation security. She served as advisory board member for ICTBIG'16, CSIBIG'14, Technical Program Committee Member ICRITO'14, ICRITO'16, Organizing Committee Member ICERDS'18, and Conference Chair NCCOCE'16, NCCOCE'17.