[10] Dilip Singh Sisodia, Vijay Khandal and Riya Singhal, "Fast prediction of web user browsing behaviours using most interesting patterns", Journal of Information Science, SAGE, Vol.44(1), 2018, pp.74-90.
[11] Marlina Abdul Latib, Saiful Adli Ismail, et.al., "Analysing Log Files For Web Intrusion Investigation Using Hadoop", ICSIE, Association for Computing Machinery, 2018, pp.12-21.
[12] Alexandra L'heureux, Katarina Grolinger, "Machine Learning With Big Data: Challenges and Approaches", Vol.5, 2017, 7776-7797.
[13] Dimitrios Sisiaridis and Olivier Markowitch, "Feature Extraction and Feature Selection: Reducing Data Complexity with Apache Spark", International Journal of Network Security & Its Applications (IJNSA) Vol.9, No.6, 2017, pp.39-51.
[14] Pinjia He, Jieming Zhu, et.al., "Towards Automated Log Parsing for Large-Scale Log Data Analysis", Journal of Latex Class Files, 2017, pp.01-14.
[15] Alex Liu, "Apache Spark Machine Learning Blue Prints", PACKT Publishing,2016.
[16] David Stodder, "Improving Data Preparation for Business Analytics", tdwi publications, 2016.
[17] Ilias Mavridis, Eleni Karatza, "Performance evaluation of cloud-based log Þle analysis with Apache Hadoop and Apache Spark", The Journal of Systems & Software, 2016.
[18] Salman Salloum, Ruslan Dautov, "Big data analytics on Apache Spark", Int J Data Sci Anal, Springer, 2016, pp.145-164.
[19] Salvador García, Sergio Ramírez-Gallego, et.al., "Big data preprocessing: methods and Prospects", Big Data Analytics, Vol. 1:9, 2016.
[20] Abdul Ghaffar Shoro & Tariq Rahim Soomro, "Big Data Analysis: Ap Spark Perspective", Global Journal of Computer Science and Technology:Software & Data Engineering, Vol. 15, Issue.1, 2015, pp.01-09.
[21] Andoena Balla, Athena Stassopoulou and Marios D. Dikaiakos, "Real-time Web Crawler Detection", 18th International Conference on Telecommunications, IEEE xplore, 2015, pp.01-05.
[22] Dr. Maheswara Rao V.V.R., N. Silpa, "A Comprehensive Study on Potential Research Opportunities of Big Data Analytics to Leverage The Transformation In Various Key Domains", International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), Vol. 5, No.5, 2015, pp.01-18.
[23] Federico Castanedo, "Data Preparation in the Big Data Era – Best Practices for Data Integration", O'Reilly Publication, 2015.
[24] Matei Zaharia, Reynold S. Xin, "Apache Spark: A Unified Engine for Big Data Processing", Communications of the ACM, Vol. 59, No. 11, 2015, pp.56-65.
[25] Mitali Srivastava, Rakhi Garg, P. K. Mishra, "Analysis of Data Extraction and Data Cleaning in Web Usage Mining", ICARCSET 2015, ACM, 2015, pp.01-06.
[26] Peng Huang, Dehua Chen, Jiajin Le, "An Improved Referrer-Based Session Identification Algorithm Using MapReduce", Ninth International Conference on Natural Computation (ICNC), IEEE, 2015, pp.1072-1076.
[27] Sara Landset, Taghi M. Khoshgoftaar, "A survey of open source tools for machine learning with big data in the Hadoop ecosystem", Journal Big Data, Journal of Big Data, Vol2.24, 2015, pp.01-36.
[28] Sonali Agarwal, Bakshi Rohit Prasad, "High Speed Streaming Data Analysis of Web Generated Log Streams", IEEE 10th International Conference on Industrial and Information Systems, ICIIS 2015, 2015, pp.01-06.
[29] Zuhair Khayyat, Ihab F. Ilyas, et.al., "BigDansing: A System for Big Data Cleansing", SIGMOD'15, ACM, 2015, pp.01-17.
[30] You Joung Ham, Hyung-Woo Lee, "Big Data Preprocessing Mechanism for Analytics of Mobile Web Log", Int. J. Advance Soft Compu. Appl, SCRG Publication, Vol. 6, No. 1, 2014, pp.01-18.
[31] Dusan Stevanovic, Aijun AN, and Natalija Vlajic, "Detecting Web Crawlers from Web Server Access Logs with Data Mining Classifiers", Springer-Verlag Berlin Heidelberg 2011, 2011, pp.483-489.
[32] Maheswara Rao V.V.R., Dr. V. Valli Kumari, Dr. KVSVN Raju "An Intelligent System for Web Usage Data Preprocessing" presented in The First International Conference on Computer Science and Information Technology - CCSIT-2011, January 2 - 4, 2011. Bangalore, India, Springer LNCS-CCIS, ISSN: 1865-0929, ISBN: 978-3-642-17856-6, Vol. 131, Part 1, pp. 481-490, 2011.
[33] Maheswara Rao V.V.R., Dr. V. Valli Kumari, Dr. KVSVN Raju "Study of Visitor Behavior by Web Usage Mining" presented in International Conference on Recent Trends in Business Administration and Information Processing - BAIP 2010, India, and proceedings published by Springer LNCS-CCIS, ISSN: 1865-0929, ISBN: 978-3-642-12213-2, Vol. 70, pp. 181-187, 2010.
[34] Pablo E. Roman, Robert F. Dell, and Juan D. Vel´asquez, "Advanced Techniques in Web Data Pre-processing and Cleaning", Advanced Techniques in Web Intelligence, Springer, 2010, pp.19-48.
[35] Jie Zhang and Ali. A. Ghorbani, "The Reconstruction of User Sessions from a Server Log Using Improved Time-oriented Heuristics", Proceedings of the 2nd Annual Conference on Communication Networks and Services Research (CNSR'04), IEEE, 2004, pp.01-08.
[36] Fang Yuan, Li-Juan Wang, Ge Yu, "Study On Data Preprocessing Algorithm In Web Log Mining", Proceedings of the Second International Conference on Machine Learning and Cybernetics, IEEE, 2003, pp.28-32

Authors Profile

N Silpa, is currently pursuing her Ph.D. in Computer Science and Engineering at Centurion University of Technology and Management (CUTM), Odissa, India. She has 10 years of teaching experience and 5 years of Research Experience. Her Research interests include Data Mining, Web Mining, Big Data Analytics, Text Mining, Data Science, Artificial Intelligence and Machine Learning.

Dr. V.V.R. Maheswara Rao, is a leading Researcher & Academician in Computer Science & Engineering and holds Ph.D. degree. He is currently working as a Professor in the Dept of Computer Science & Engineering at Shri Vishnu Engineering College for Women (A), Andhra Pradesh, India. He is actively involved and successfully implemented three projects funded by DST. He has 31 research papers, six of which are Scopus-indexed and four of which are Web of Science-indexed. He has 23 years of experience that include 6 years of Industry experience, 17 years of Teaching experience and 13 years of Research experience. His Research interests include Data Mining, Web Mining, Big Data Analytics, Data Science, Artificial Intelligence and Machine Learning.